Daily report 20-05-2020

Analysis and prediction of COVID-19 for EU-EFTA-UK and other countries

Situation report 67

Contact: clara.prats@upc.edu

With the financial support of:

Foreword

The present report aims to provide a comprehensive picture of the **pandemic situation of COVID-19** in the EU countries, and to be able to foresee the situation in the next coming days.

We employ an **empirical model**, verified with the evolution of the number of confirmed cases in previous countries where the epidemic is close to conclude, including all provinces of China. The model does not pretend to interpret the causes of the evolution of the cases but to permit the **evaluation of the quality of control measures made in each state** and a **short-term prediction of trends**. Note, however, that the effects of the measures' control that start on a given day are not observed until approximately 7-10 days later.

The model and predictions are based on two parameters that are daily fitted to available data:

- ✓ *a*: the velocity at which spreading specific rate slows down; the higher the value, the better the control.
- ✓ *K*: the final number of expected cumulated cases, which cannot be evaluated at the initial stages because growth is still exponential.

We show an individual report with 8 graphs and a table with the **short-term predictions** for different countries and regions. We are adjusting the model to **countries and regions** with at least 4 days with more than 100 confirmed cases and a current load over 200 cases. The **predicted period** of a country depends on the number of datapoints over this 100 cases threshold, and is of 5 days for those that have reported more than 100 cumulated cases for 10 consecutive days or more. For short-term predictions, we assign higher weight to last 3 points in the fittings, so that changes are rapidly captured by the model. The whole methodology employed in the inform is explained in the last pages of this document.

In addition to the individual reports, the reader will find an initial dashboard with a brief analysis of the situation in EU-EFTA-UK countries, some summary figures and tables as well as **long-term predictions** for some of them, when possible. These long-term predictions are evaluated without different weights to data-points. We also discuss a specific issue every day.

Martí Català
Pere-Joan Cardona, PhD
Comparative Medicine and Bioimage Centre of Catalonia; Institute for Health Science Research Germans Trias i Pujol

Raquel Pérez-Arnal
Dario Garcia-Gasulla
Sergio Alvarez-Napagao
Barcelona Supercomputing Center

Clara Prats, PhD
Sergio Alonso, PhD
Enric Álvarez, PhD
Miquel Marchena
David Conesa
Daniel López, PhD
Computational Biology and Complex Systems; Universitat Politècnica de Catalunya - BarcelonaTech

With the collaboration of: Guillem Álvarez, Oriol Bertomeu, Laura Dot, Lavínia Hriscu, Helena Kirchner, Daniel Molinuevo, Pablo Palacios, Sergi Pradas, David Rovira, Xavier Simó, Tomás Urdiales

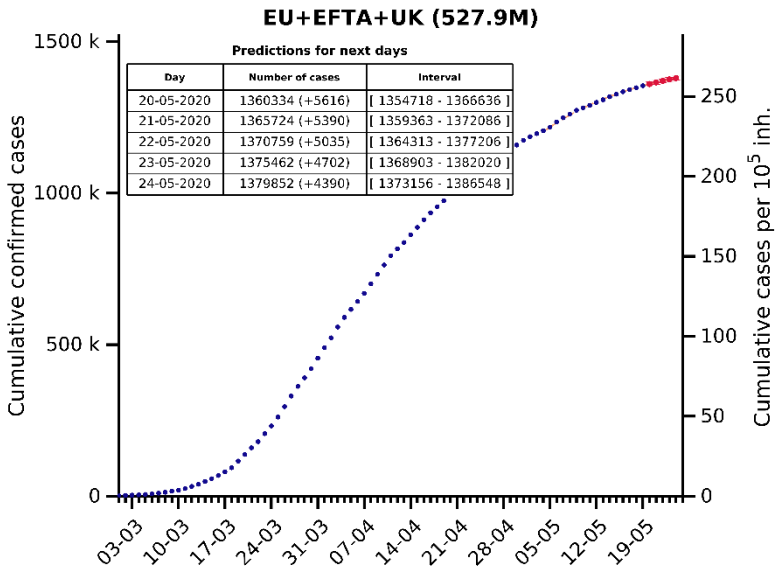
PJC and MC received funding from "la Caixa" Foundation (ID 100010434), under agreement LCF/PR/GN17/50300003; CP, DL, SA, MC, received funding from Ministerio de Ciencia, Innovación y Universidades and FEDER, with the project PGC2018-095456-B-I00;

Disclaimer: These reports have been written by declared authors, who fully assume their content. They are submitted daily to the European Commission, but this body does not necessarily share their analyses, discussions and conclusions.

(0) Executive summary – Dashboard

Global EU+EFTA+UK trends and needs

Predicting the behavior of the epidemic by country and region has been relatively feasible, so far. Gompertz's model has described quite well the observed behavior. Similar results can be achieved with other empirical models, and even with properly fitting of SIR or SEIR-type models. Now, step by step, **all countries are entering the final stage where the behavior of the epidemic presents particular characteristics**.



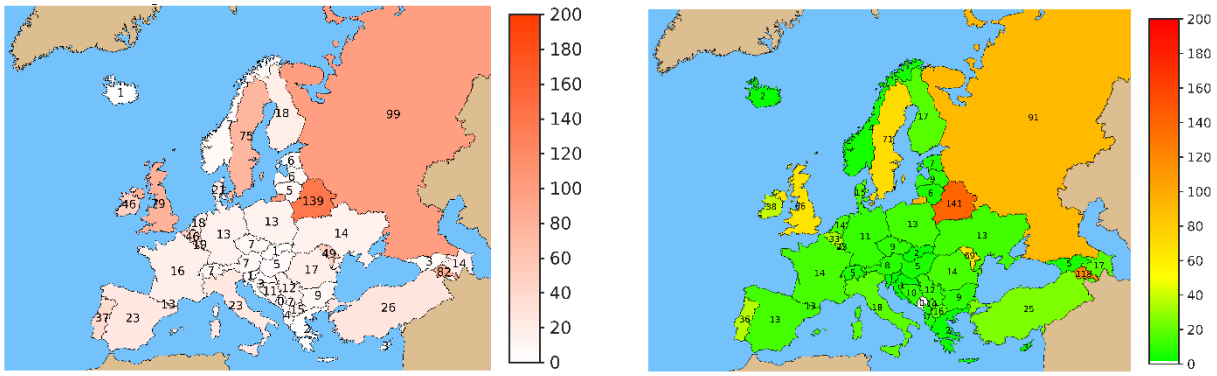
We can take as example the dynamics in three countries that have already reached this stage: Austria, Luxembourg and Iceland. Looking at their plots of daily new cases, we see **how its tail lengthens**. We observe, for example, how Austria presents an approximately constant number of daily cases since May 2, with an average of 44 daily. It is really difficult to predict whether this value will be lowered or whether it will remain approximately constant for a long time. In Luxembour we observe a similar behavior, but in this case the level is at around 8 cases per day since May 9. Finally, in Iceland we find that they have managed reduice the number of daily cases down to zero. This is the challenge for all countries, that should reach one of these situations gradually.

The analysis is focused on **comparing two mobility indices: one provided by Facebook Data for Good and the other supplied by Google**. Those indicators can be especially useful on the deconfinement processes, since can anticipate an increase in probability of a secondary outbreak. This analysis has been elaborated together with researchers from the Barcelona Supercomputing Center.

Trends for specific countries

Austria and **Luxembourg** show a ρ_7 around 1. This is characteristic of the mentioned long tail, where the number of new cases remain constant at a certain low level. A constant number of new cases means that each new infected is infecting another person. If the number of active cases is low, as is the case of these countries, then the risk remains low (EPG_{REP} of 8 and 24, respectively).

The map in the left shows current $\mathbf{A_{14}}$. The map in the right shows current **EPG**.



Situation and trends per country

Table of current situation in EU countries. Colour scale is relative except when indicated, this means that it is applied independently to each column, and distinguishes best (green) form worst (red) situations according to each of the variables. Last column (**EPG_{EST}**) indicates EPG assessed with **estimated real 14-day attack rate** (see report from 22/04 for details). **EPG_{REP}** is calculated with **data reported by countries**. EPG_{REP} and EPG_{EST} **cannot be compared between them** because scales are different, but can be independently used for estimating risk of countries according to reported or estimated real situation, respectively.

	Reported data						Indexes		
Country	Cumulative cases	Attack rate /10^5 inh.	Cumulative deaths	Mortality /10^5 inh.	Active cases (last 14 days)	14-day attack rate /10^5 inh.	ρ_7 [1]	EPG_{REP} [2]	EPG_{EST} [3]
United Kingdom	248,818	374.5	35,341	53.2	53,828	81.0	0.83	67	969
Spain	232,037	500.6	27,778	59.9	10,913	23.5	0.58	14	162
Italy	226,699	381.5	32,169	54.1	13,686	23.0	0.78	18	252
Germany	176,007	214.9	8,090	9.9	11,110	13.6	0.81	11	52
France	143,427	221.6	28,022	43.3	10,460	16.2	0.87	14	256
Belgium	55,791	491.2	9,108	80.2	5,282	46.5	0.73	34	556
Netherlands	44,249	260.5	5,715	33.6	3,162	18.6	0.74	14	179
Sweden	30,799	313.1	3,743	38.0	7,583	77.1	0.94	73	960
Switzerland	30,535	356.3	1,613	18.8	609	7.1	0.74	5	28
Portugal	29,432	283.8	1,247	12.0	3,730	36.0	0.97	35	152
Ireland	24,251	513.1	1,561	33.0	2,268	48.0	0.83	40	257
Poland	19,268	50.4	948	2.5	4,837	12.7	1.03	13	76
Romania	17,191	86.9	1,126	5.7	3,354	17.0	0.80	14	99
Austria	16,257	186.6	632	7.3	671	7.7	1.07	8	32
Denmark	11,044	193.3	551	9.6	1,223	21.4	0.70	15	75
Czech Republic	8,647	81.5	302	2.8	751	7.1	1.35	10	35
Norway	8,257	153.8	233	4.3	354	6.6	0.61	4	12
Finland	6,399	116.3	301	5.5	987	17.9	0.93	17	82
Luxembourg	3,958	687.2	109	18.9	118	20.5	1.19	24	63
Hungary	3,598	36.9	470	4.8	487	5.0	1.09	5	78
Greece	2,840	25.4	165	1.5	198	1.8	1.17	2	13
Bulgaria	2,292	32.1	116	1.6	603	8.5	1.09	9	52
Croatia	2,232	53.0	96	2.3	120	2.8	0.32	1	NA
Iceland	1,802	494.7	10	2.7	3	0.8	0.29	0	NA
Estonia	1,791	136.5	64	4.9	80	6.1	1.13	7	NA
Lithuania	1,562	53.7	60	2.1	139	4.8	1.08	5	NA
Slovakia	1,495	27.5	28	0.5	74	1.4	1.35	2	NA
Slovenia	1,467	70.6	104	5.0	22	1.1	0.50	1	4
Latvia	1,012	51.3	21	1.1	116	5.9	1.52	9	NA
Cyprus	918	78.5	17	1.5	40	3.4	0.77	3	NA
Malta	560	130.5	6	1.4	78	18.2	NA	NA	NA
Liechtenstein	83	215.3	1	2.6	0	0.0	NA	NA	NA
	Scale								
	Worst	Worst	Worst	Worst	Worst	Worst	2.0	200	2000
	Best	Best	Best	Best	Best	Best	0.0	0	0

[1] ρ_3 is the average of 7 consecutive ρ, but can still fluctuate. [2,3] EPG stands for Effective Growth Potential. EPG_{REP} is obtained by multiplying attack rate of last 14 days per 10^5 inhabitants (i.e. density of cases) by ρ_7 (a value related with effective reproduction number and that, therefore, determines the dynamics for subsequent days). EPG_{EST} is obtained by multiplying estimated real attack rate of last 14 days per 10^5 inhabitants by ρ_7.

Highlights for countries with highest number of reported cases

- ✓ Spain is revising historical series of other regions with minor variations, including a one-day peak of 2,700 cases in Catalunya (so, ρ_7 not reliable). Data by region have not been updated today. Therefore, they are not included in the detailed report.
- ✓ UK could reach the level of 2,000 daily new cases in a few days. Italy and Germany are at the level of 500, while France and Spain seem to be at 300-200 daily new cases.
- ✓ UK, Italy, Germany and France have a ρ_7 of 0.8-0.9.

Time indicators by country

This table summarizes a few time indicators for each country: time since 50 cases were reported, time interval between an attack rate of $1/10^5$ inhabitants and an attack rate of $10/10^5$ inhabitants, and time interval between attack rates of 10 to 100 per 10^5 inhabitants (only for countries that have overtaken this threshold).

Countries	Days since the first 100 cases	Time interval between 1 and 10 cases / 10^5 inh. (days)	Time interval between 10 and 100 cases / 10^5 inh. (days)
Italy	87	11	16
Germany	81	12	17
France	80	10	20
Spain	80	8	12
United Kingdom	76	10	12
Belgium	75	11	14
Netherlands	75	11	20
Sweden	75	10	28
Norway	75	2	7
Switzerland	75	8	11
Austria	73	10	14
Denmark	72	4	30
Czech Republic	69	11	NA
Finland	69	12	46
Greece	69	18	NA
Iceland	69	5	15
Portugal	68	9	15
Slovenia	68	6	NA
Estonia	67	5	30
Ireland	67	8	18
Poland	67	17	NA
Romania	67	15	NA
Luxembourg	64	6	7
Slovakia	63	24	NA
Bulgaria	62	30	NA
Croatia	62	12	NA
Hungary	61	20	NA
Latvia	61	12	NA
Lithuania	60	9	NA
Malta	59	9	35
Cyprus	58	12	NA

Analysis: Assessments on mobility data on confinement and deconfinement (I).

The report yesterday showed that the level of available daily tests per 100,000 people together with the number of estimated active cases gives a very good sense of the ability of a country or region to implement a test and trace program. However, the analysis required that the number of close contacts of possible active cases with people with no kinship in public spaces is small. Most European countries have policies in place that forbid the gathering of a large number of people and, in some countries, the number of restriction measures is large. Most countries are in the process of lifting these restrictions gradually. It is thus important to have measure of how the mobility is returning back to normal.

We must stress the fact that **while people might have the same level of mobility as before and lead to the same number of contacts, these contacts might be way less problematic**. If masks, increase in physical distance and hygienic measures (i.e., soap and hydroalcoholic solutions) are commonly used systematically to protect others and oneself, the same number of contacts will lead to less infections. The **mobility measures**, if properly assessed and treated, give us an **upper level, or worst-case scenario, of the situation**. They should stablish if the number of contacts is reaching pre-pandemic levels not if those contacts produce the same level of infections.

It is thus important to develop a **reliable measure of these contacts**. We state here that the relevant measure for these contacts is not so much the number of trips between cities but purely the fact of leaving home and connecting to other people at work, shop, leisure, etc. The connectivity between cities and towns is important in order to track the possible propagation of an outbreak, but implementing control measures for long trains/buses and planes is relatively easy and we expect all countries to use them. Therefore, we consider confinement and deconfinement measures the relevant ones.

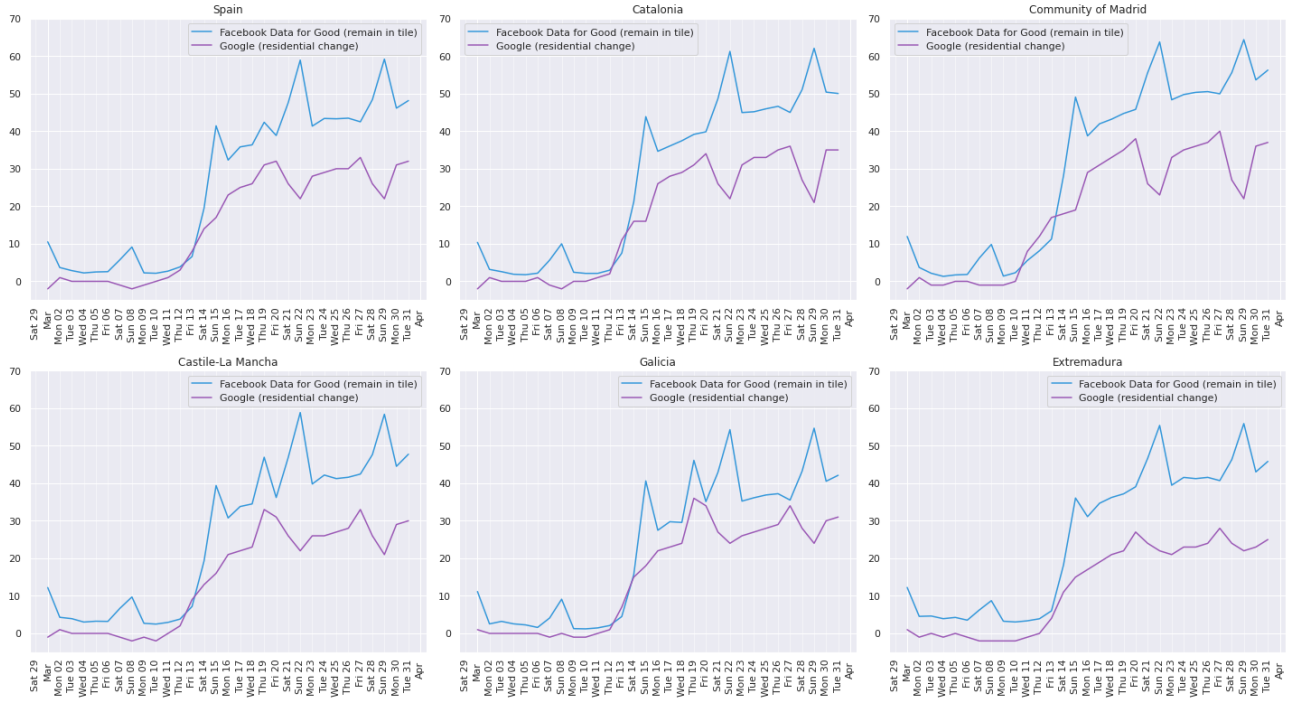
There are, right now, two reliable **measurements of confinement** provided by **Facebook Data for Good** (FDG) and by **Google** (GGL). **FDG** provides the **fraction of people who remain in the same area** (roughly 500 by 500 meters) during the whole aggregated by level 3 (provinces in Spain). On the other hand, **GGL** measures the **increase in relative time, compared with a baseline, that people remain in their own residence** (with an accuracy of around 100m in urban areas[1]), with residence defined as the spot where they spend most of the time, especially during the night[2]. A key difference between FDG and GGL is that the later reports aggregated by level 2, which corresponds to a Spanish Autonomous Community (AC) or an Italian region. Each AC is formed by one or more provinces. In order to compare GGL and FDG data, we need to aggregate FDG up to level 2 using population of each province.

Other confinement indices and measurements have shown problems or have no clear definitions and have been disregarded. The key point of this report is to show that both mobility measurements are perfectly correlated in a country like Spain, which is representative of the average European country in terms of mobile coverage and fraction of users. Given that we are active partners of FDG we can guarantee a relative representative population (the 35-65 age bracket is over-represented in these indicators) and indicate that most European countries have similar behavior. Given that GGL does not inform of most of its processes, we take just them as a black box index that we use to have a reference, but where we cannot guarantee its representativeness.

Spain took confinement measures on the 15th of March, and we have shown that FDG index showed a large increase in the confinement's measurement just after this order was issued, while Italy showed a more gradual increase. We are happy to observe that the GGL measurement increases equally sharp, as illustrated

1 https://www.tandfonline.com/doi/pdf/10.1080/20961790.2018.1509187
2 https://arxiv.org/pdf/2004.04145.pdf

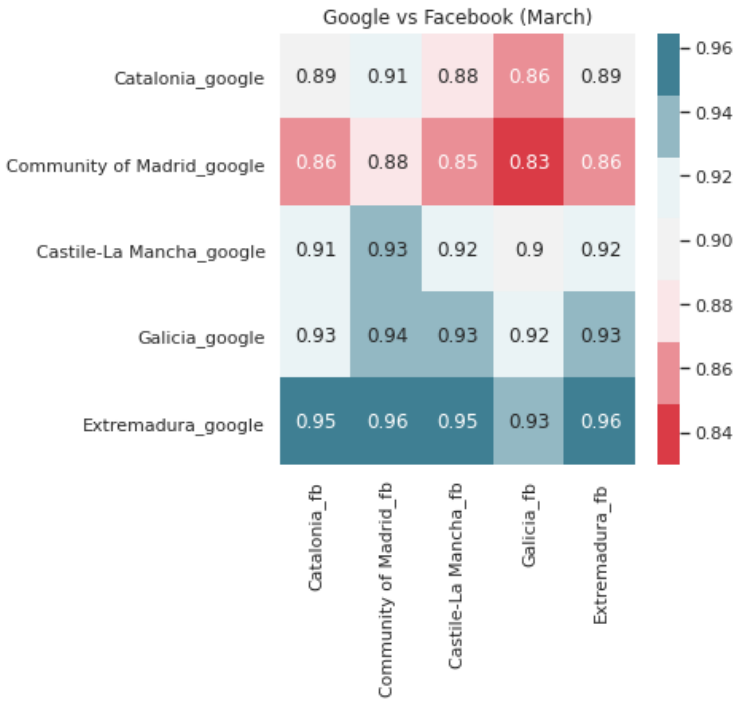
by the next figure. In the plots, the index for FDG has a slight modification since the number of people under normal mobility that remain in tile is roughly 10%. We have subtracted this value to the FDG series so that both indexes are easier to compare.



We show above the evolution of both indexes during March for the whole country and for Autonomous Communities with different densities and populations like Madrid, Catalunya, Castilla-La Mancha, Galicia, and Extremadura. Madrid and Catalunya are examples of regions with high incidence and very large densities (830 and 240 people/km^2). Castilla-La Mancha is an example of AC with the highest incidence and very low density (25 people/km^2): the reason for this high incidence and low density is its direct connection with Madrid. Galicia and Extremadura represent areas with low density (90 and 25 people/km^2) and intermediate and low incidence, respectively, which seems related with density.

We observe that both indexes show a very important increase when the confinement measures are implemented. They are not the same because **one measures absolute ratios of the people remaining in a given area** (FDG), while the **other measures relative changes** (GGL). In addition, time scales are different for both indices, and there is a discrepancy on Sundays that will be discussed in upcoming reports. The important thing is that both correlate strongly in the same direction, pointing to common underlying measurement that can be properly assessed and, in principle, combined in a single index in order to give a unifying and correct picture of the situation.

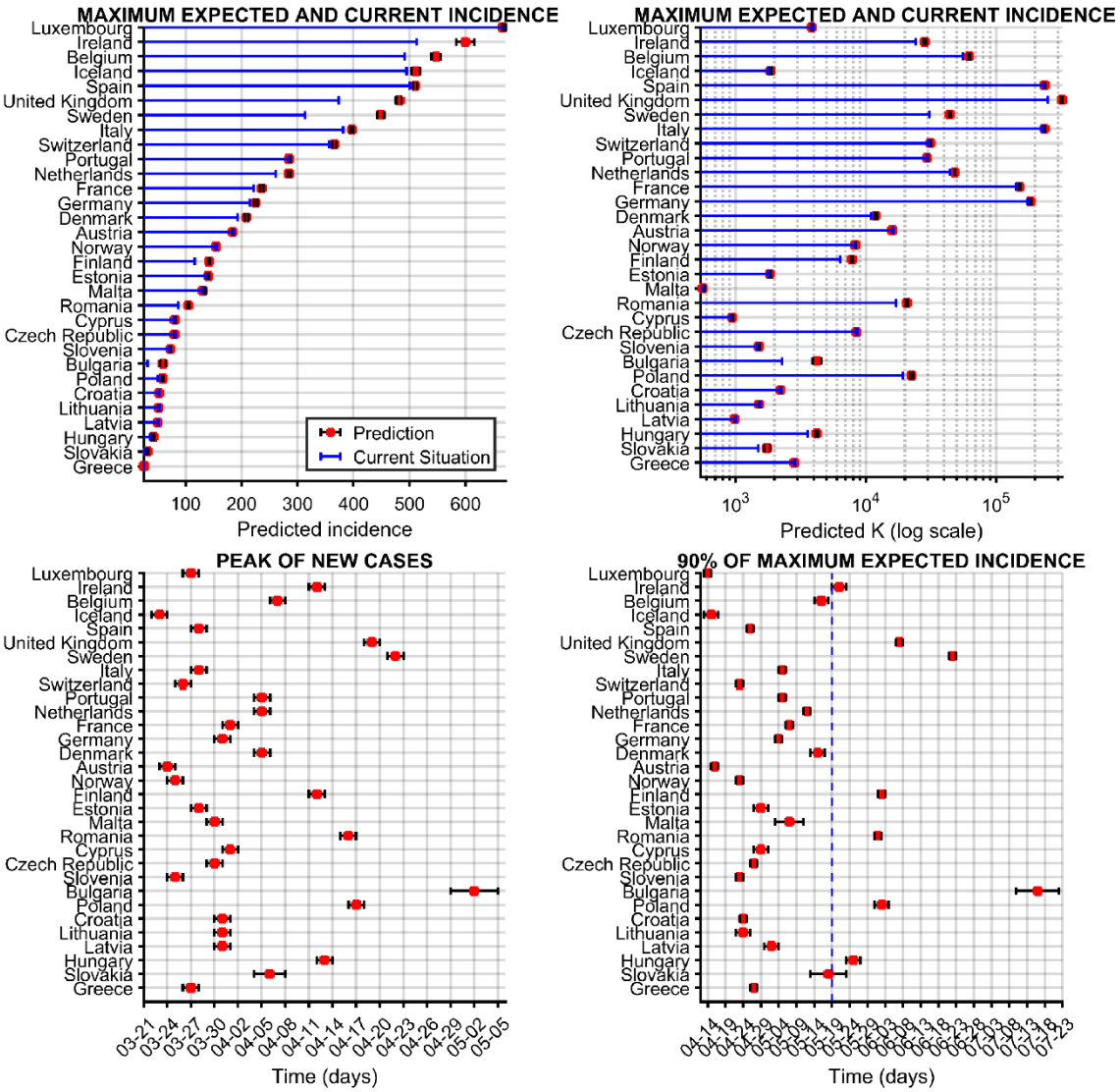
We can proceed to test whether the correlation of both indices during this period is high or not. Below, you can find a table of the correlation coefficient between both measurements and between these regions. We can check that correlations are extremely high. Both show the same information, and all communities basically go in the same direction showing a general compliance of the order and that **both measurements indicate basically the same information**.



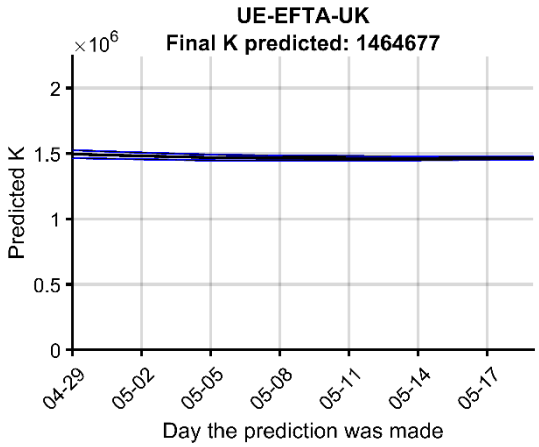
Having shown that sharp movement are perfectly correlated, we leave for the next report to study if this correlation holds one the process of confinement is stablished and slowly lifted. This will shed light regarding the proper way to combine them with the final end of generating a single mobility index.

Long-term predictions

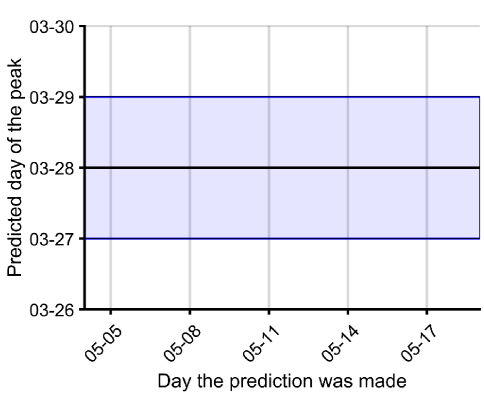
Long-term predictions, evaluated with the **whole historical series** and without weighting last 3 points. Up-left: Predictions of maximum incidences per country (total final expected attack rate per 10^5 inh.). Up-right: Predictions of maximum absolute number of cases per country (K, in log scale). Blue lines indicate current situation. Bottom-left: Time in which peak in new cases was achieved / will be achieved. Bottom-right: Time at which 90 % of K was achieved / will be achieved. Blue dotted line indicates current date.



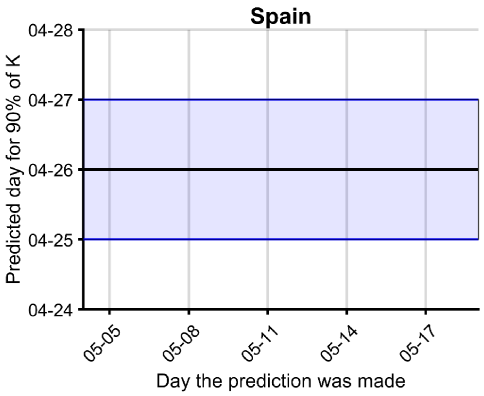
Final expected K for UE+EFTA+UK. Evolution of predicted K with time, where convergence to best estimate is seen. Last prediction is numerically shown in title.

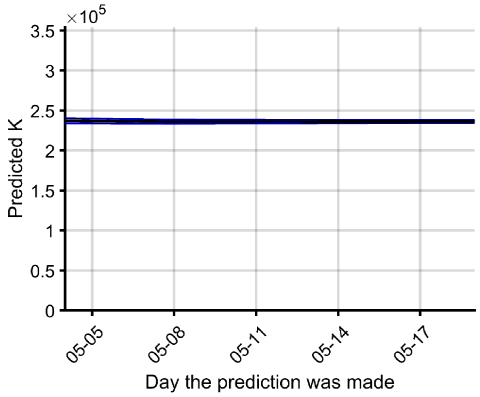


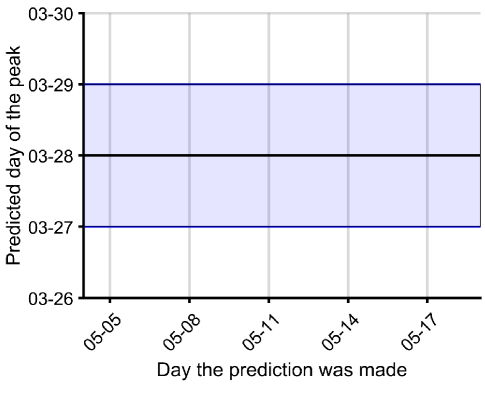
2020-05-19



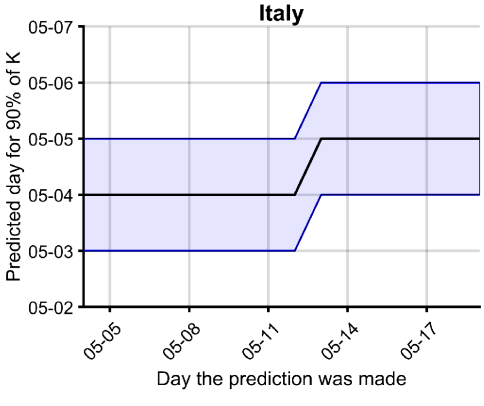
Spain

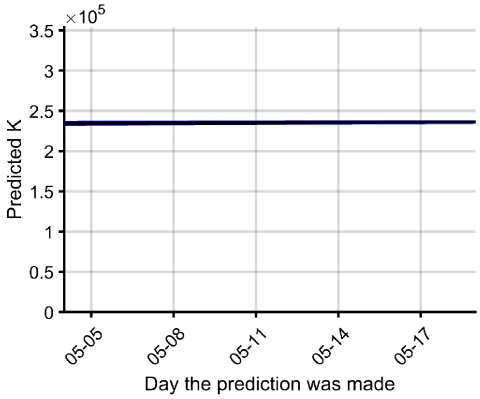


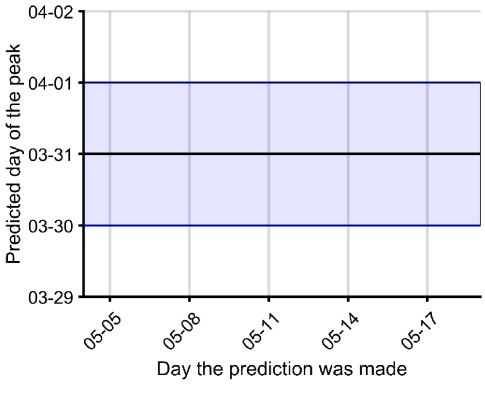




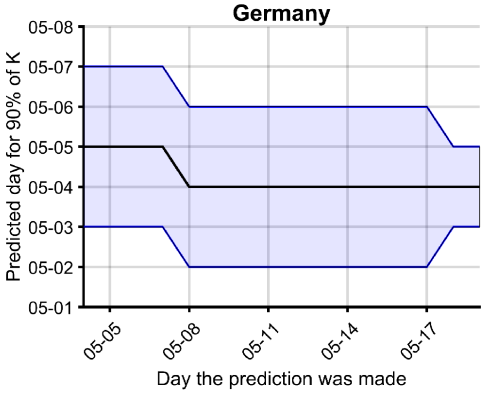
Italy

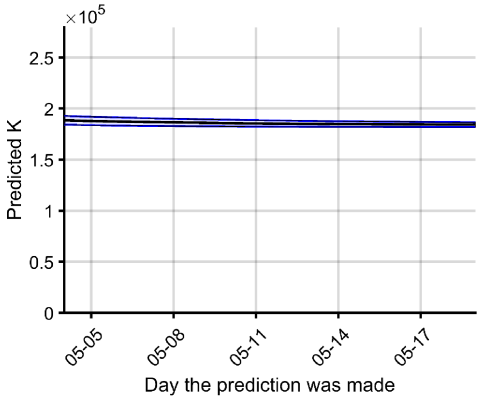


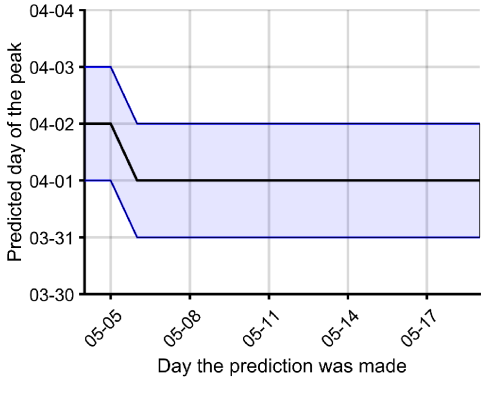




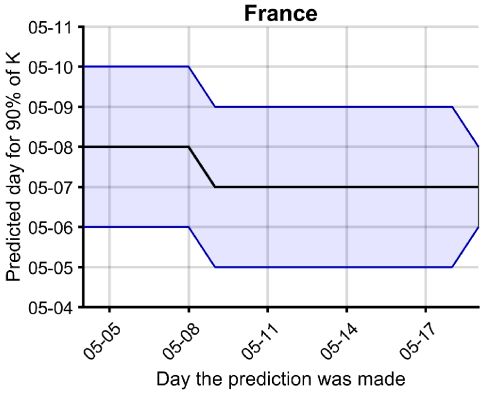
Germany

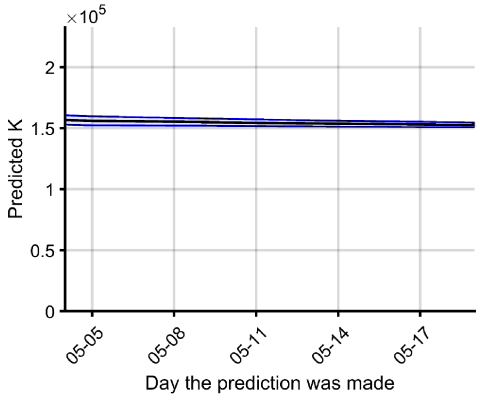


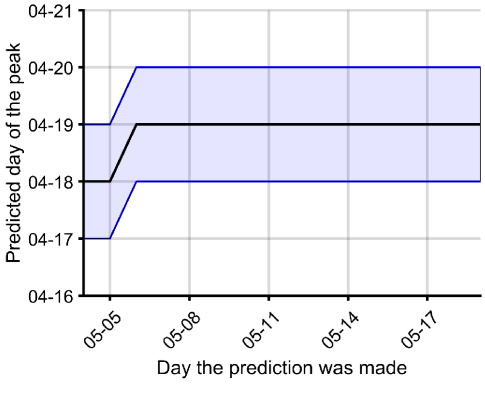




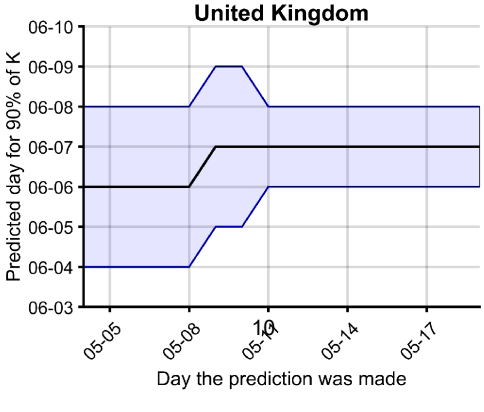
France

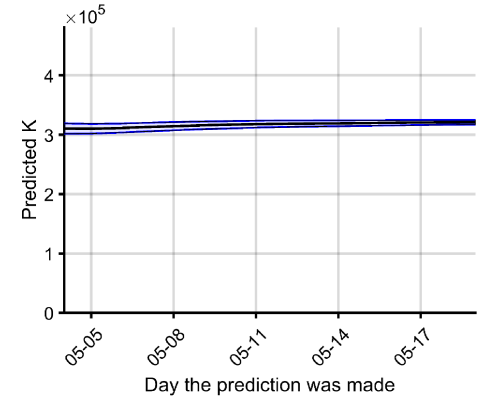




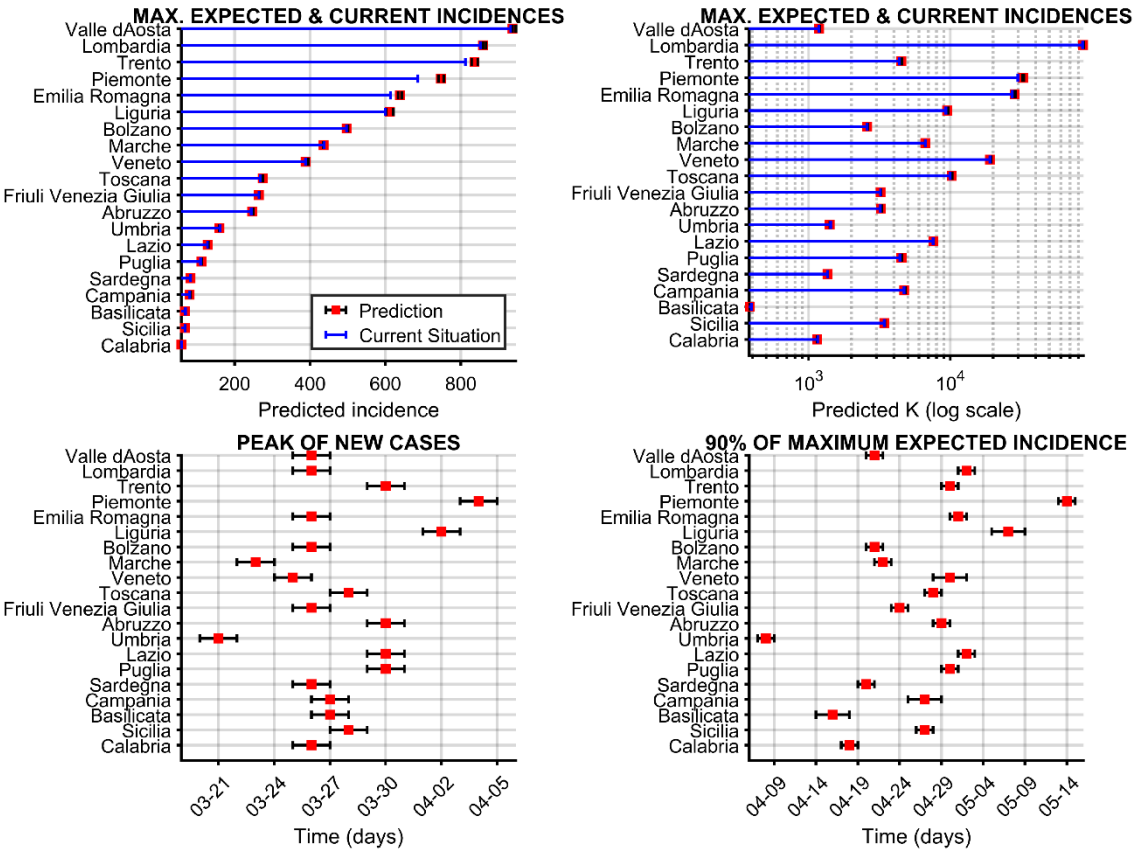


United Kingdom





Italian regions



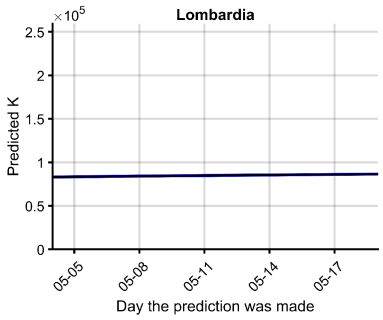
Situation and trends in Italian and Spanish regions

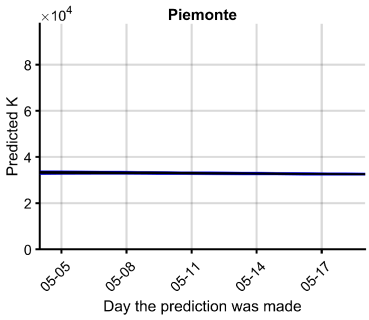
Italy

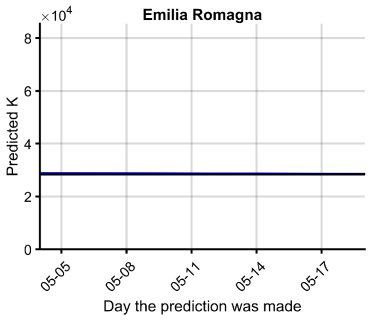
	Reported data						Indexes		
Country	Cumulative cases	Attack rate / 10^5 inh.	Cumulative deaths	Mortality / 10^5 inh.	Active cases (last 14 days)	14-day attack rate / 10^5 inh	ρ_7 [1]	EPG_{REP} [2]	EPG_{EST} [3]
Lombardia	85.775	854,2	15.662	156,0	6.406	63,8	0,76	49	901
Piemonte	29.885	686,0	3.718	85,3	1.946	44,7	0,80	36	448
Emilia Romagna	27.364	613,6	4.008	89,9	985	22,1	0,71	16	232
Veneto	19.030	387,9	1.832	37,3	551	11,2	0,73	8	81
Toscana	9.982	267,6	998	26,8	325	8,7	1,02	9	91
Liguria	9.289	599,0	1.386	89,4	738	47,6	1,10	52	790
Lazio	7.533	128,1	647	11,0	538	9,2	1,19	11	100
Marche	6.677	437,8	987	64,7	256	16,8	0,70	12	175
Campania	4.714	81,3	401	6,9	182	3,1	0,94	3	25
Puglia	4.407	109,4	478	11,9	211	5,2	0,55	3	31
Trento	4.368	407,4	455	42,4	88	8,2	1,68	14	303
Sicilia	3.411	68,2	268	5,4	130	2,6	0,77	2	16
Friuli Venezia Giulia	3.209	264,1	322	26,5	115	9,5	0,98	9	93
Abruzzo	3.205	244,4	389	29,7	158	12,0	1,20	14	179
Bolzano	2.587	2.407,9	291	270,9	44	41,0	0,65	27	59
Umbria	1.427	161,8	74	8,4	23	2,6	0,84	2	NA
Sardegna	1.355	82,6	126	7,7	36	2,2	0,60	1	12
Valle dAosta	1.175	935,5	143	113,8	29	23,1	0,87	20	242
Calabria	1.156	59,4	96	4,9	34	1,7	0,91	2	NA
Molise	422	138,1	22	7,2	118	38,6	0,39	15	NA
Basilicata	393	69,8	27	4,8	NA	NA	NA	NA	NA
Scale									
Worst	Worst	Worst	Worst	Worst	Worst	Worst	2,0	200	2000
Best	Best	Best	Best	Best	Best	Best	0,0	0	0

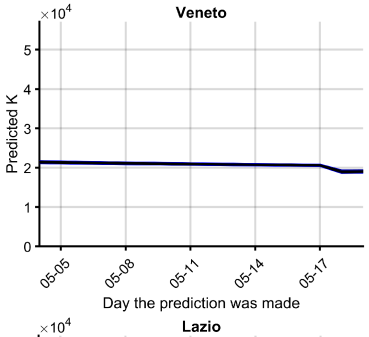
[1] ρ_7 is the average of 7 consecutive ρ, but can still fluctuate. [2,3] EPG stands for Effective Growth Potential. EPG_{REP} is obtained by multiplying attack rate of last 14 days per 10^5 inhabitants (i.e. density of cases) by ρ_7 (a value related with

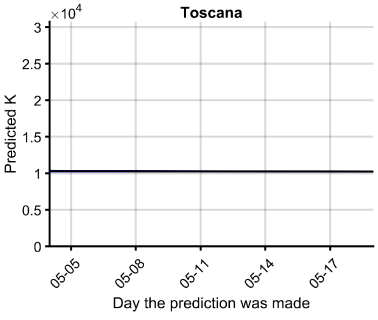
2020-05-20

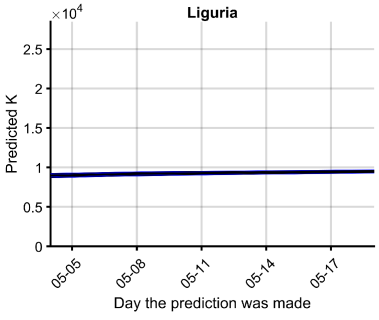

Lombardia
×10^5
Predicted K
Day the prediction was made


Piemonte
×10^4
Predicted K
Day the prediction was made

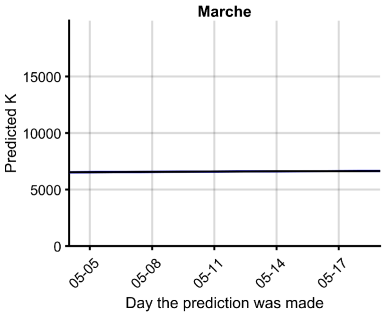

Emilia Romagna
×10^4
Predicted K
Day the prediction was made

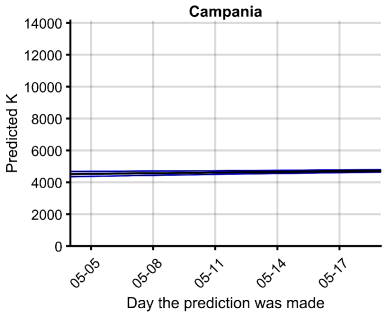

Veneto
×10^4
Predicted K
Day the prediction was made

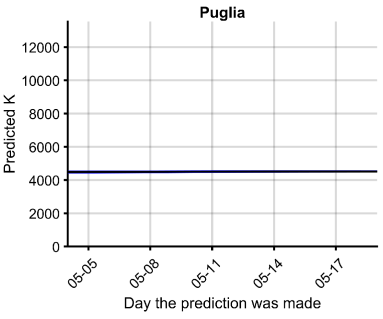

Toscana
×10^4
Predicted K
Day the prediction was made

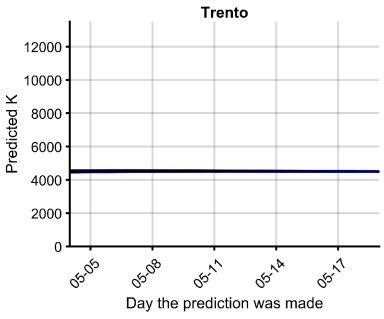

Liguria
×10^4
Predicted K
Day the prediction was made

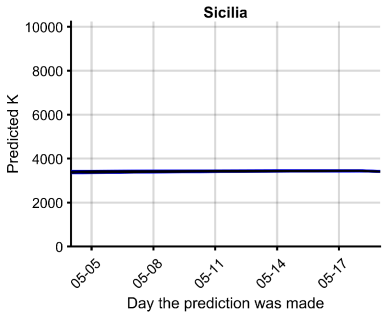

Lazio
×10^4
Predicted K
Day the prediction was made

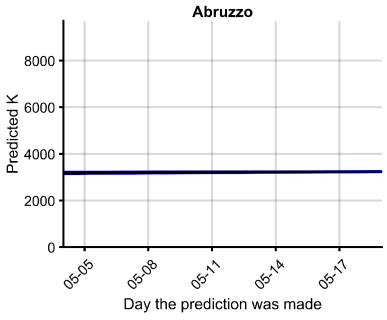

Marche
Predicted K
Day the prediction was made

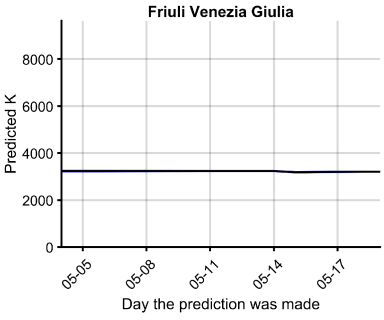

Campania
Predicted K
Day the prediction was made

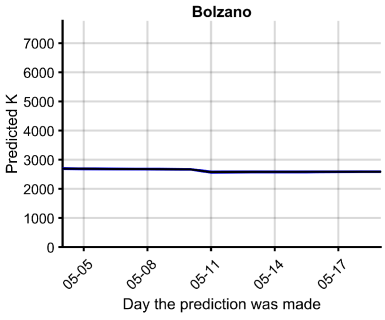

Puglia
Predicted K
Day the prediction was made

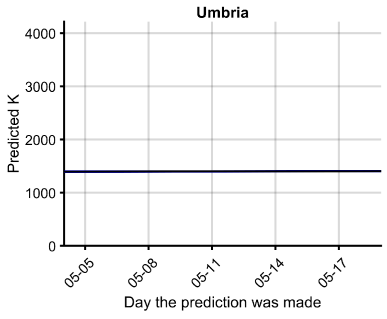

Trento
Predicted K
Day the prediction was made

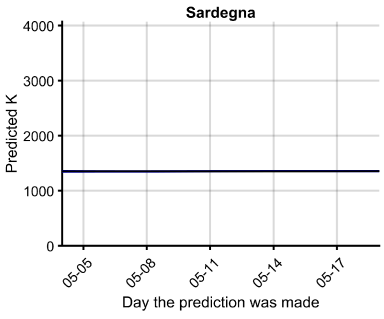

Sicilia
Predicted K
Day the prediction was made

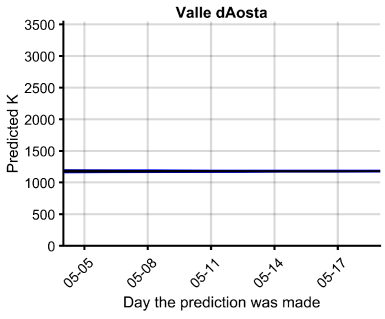

Abruzzo
Predicted K
Day the prediction was made

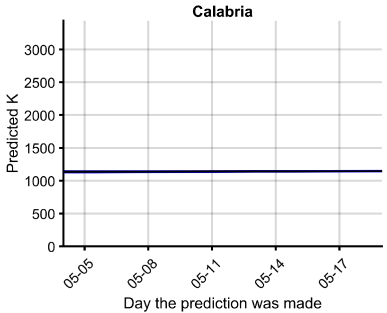

Friuli Venezia Giulia
Predicted K
Day the prediction was made

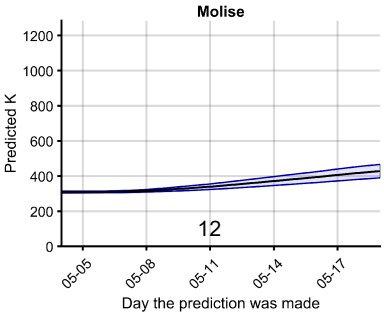

Bolzano
Predicted K
Day the prediction was made

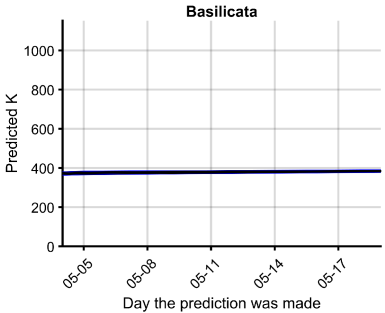

Umbria
Predicted K
Day the prediction was made


Sardegna
Predicted K
Day the prediction was made


Valle dAosta
Predicted K
Day the prediction was made

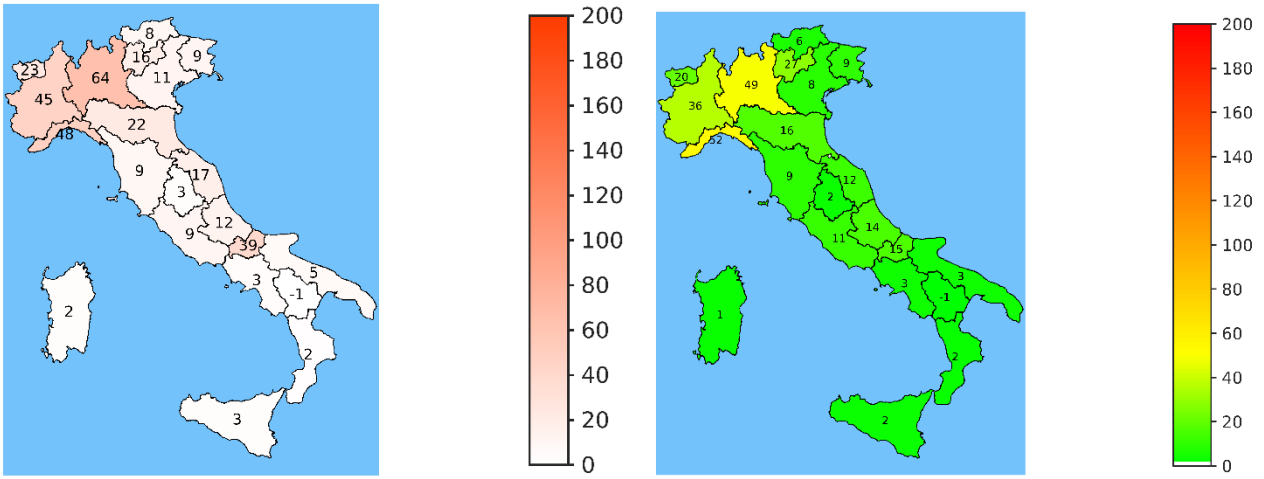

Calabria
Predicted K
Day the prediction was made


Molise
Predicted K
Day the prediction was made


Basilicata
Predicted K
Day the prediction was made

effective reproduction number and that, therefore, determines the dynamics for subsequent days). EPG_{EST} is obtained by multiplying estimated real attack rate of last 14 days per 10^5 inhabitants by ρ_7.

The **maps** in the left show current $\mathbf{A_{14}}$. The maps in the right show current **EPG**.

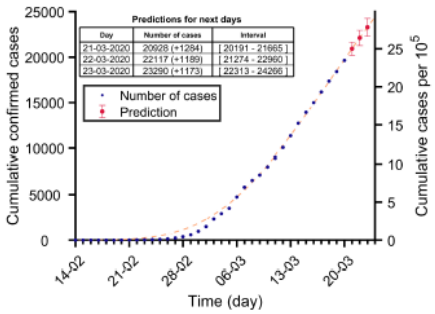


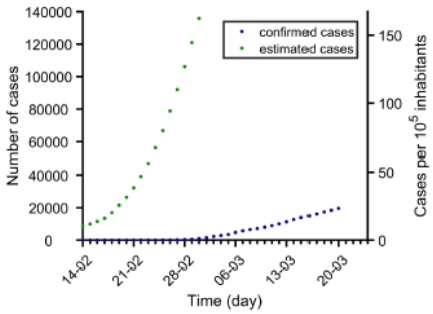
Spain: Data by region have not been updated today. Therefore, they are not included in the detailed reports.

Legend: Countries' reports details

Iran 20-03-2020. Population: 83.7M. Current cumulated incidence: $23/10^5$

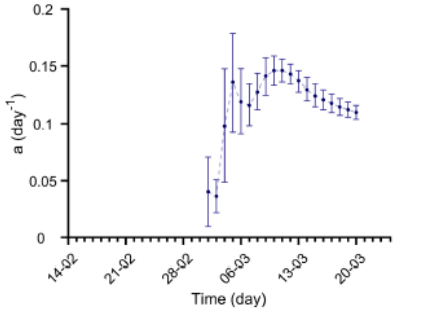
Confirmed cases: data (blue), model fitted (dashed line), predictions (red points and table)

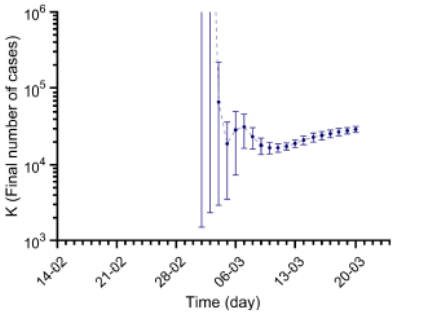




Estimated cases using death rate (see Methods)

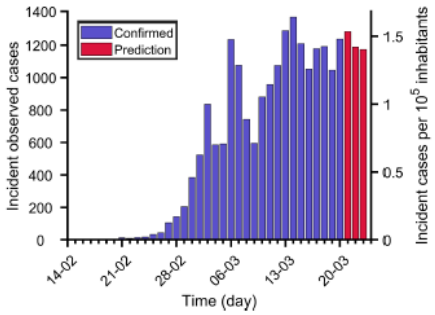
Fitted *a* value using points prior to each date

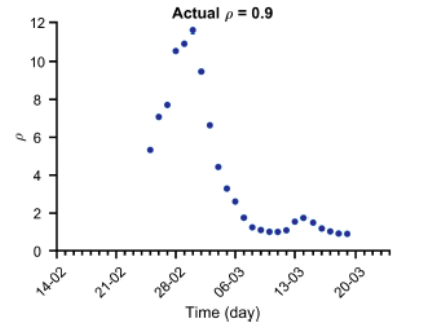




Fitted *K* value using points prior to each date

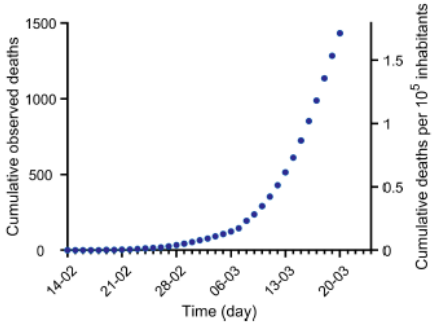
Reported and predicted new cases

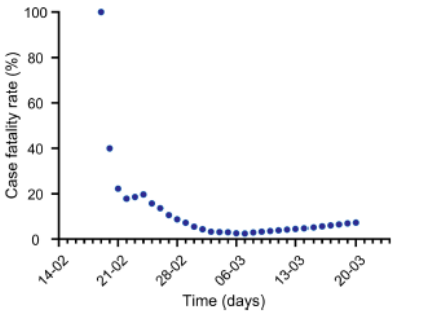




Evolution of ρ, a parameter related with Reproduction number (see Methods)

Reported deaths



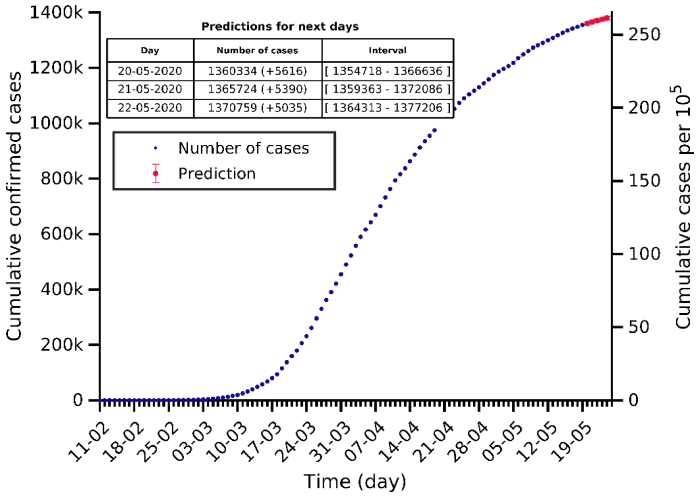


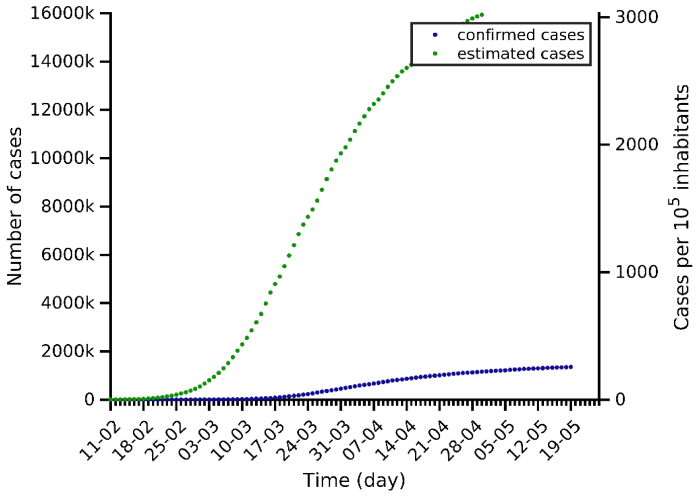
Deaths / cumulated reported cases

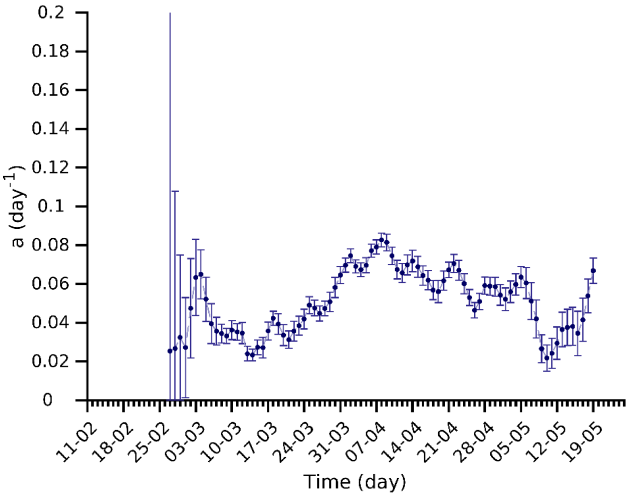
(1) Analysis and prediction of COVID-19 for EU+EFTA+UK

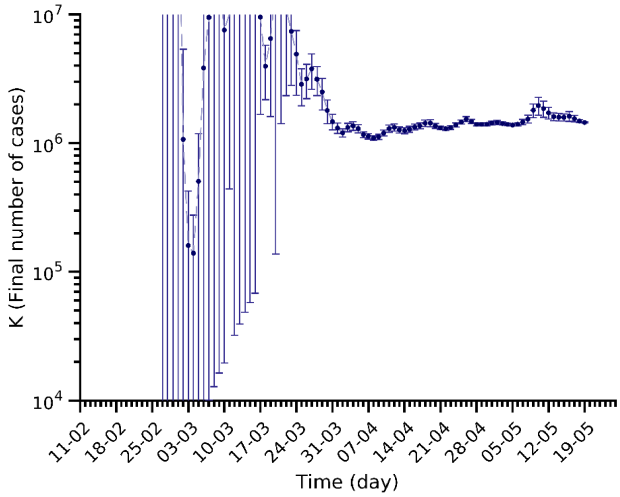
Data obtained from https://www.ecdc.europa.eu/en/geographical-distribution-2019-ncov-cases

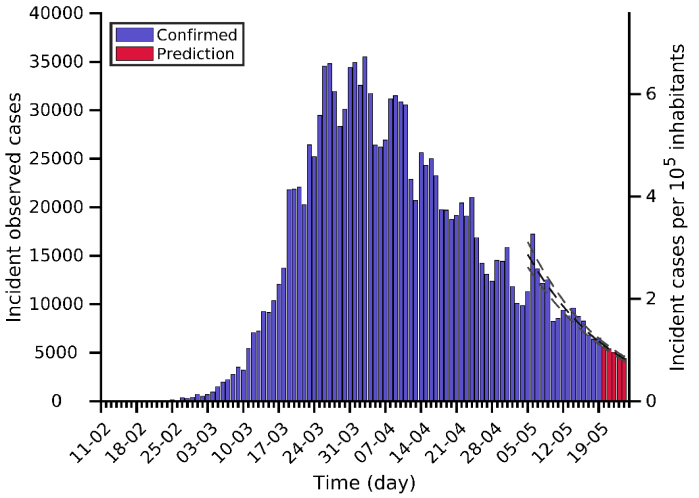
EU+EFTA+UK 19-05-2020. Population: 527.9M. Current cumulated incidence: 257/1

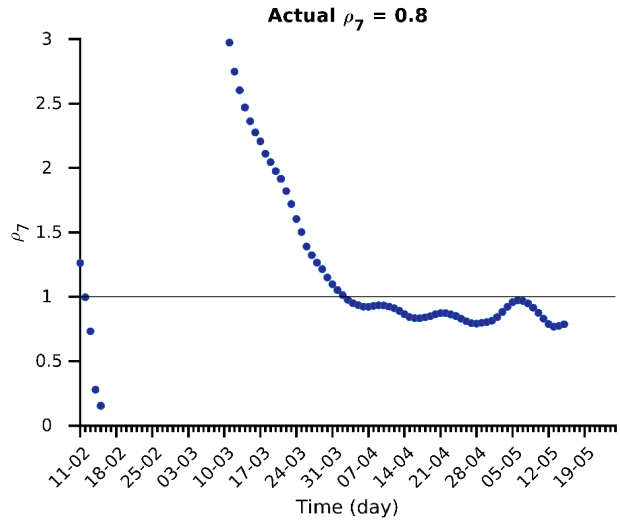


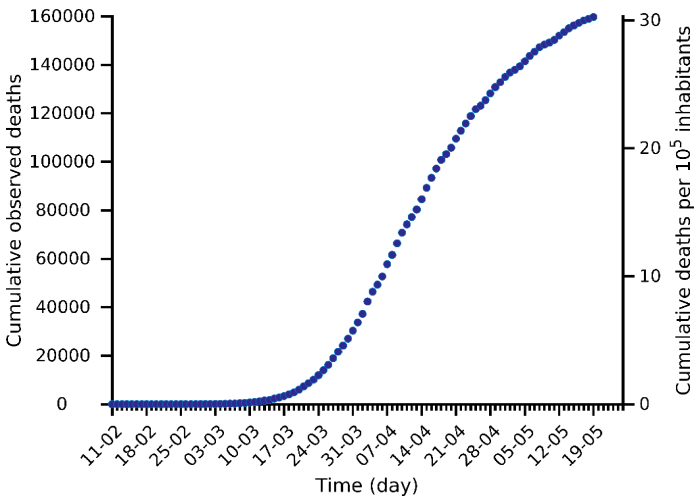


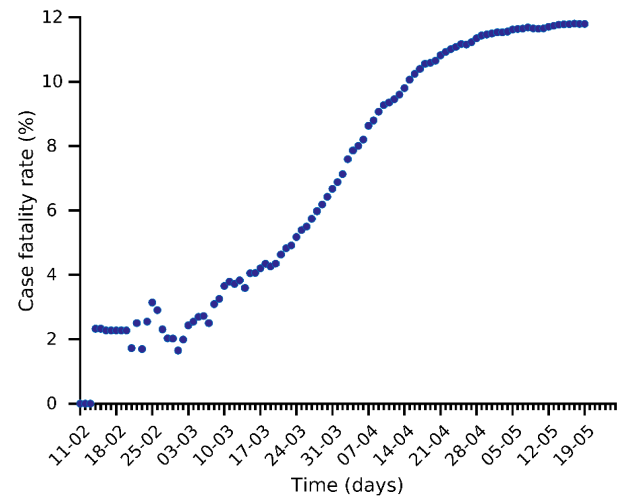




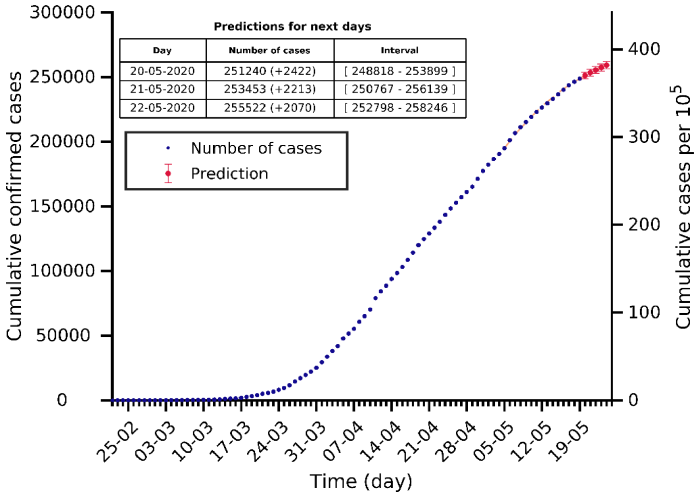


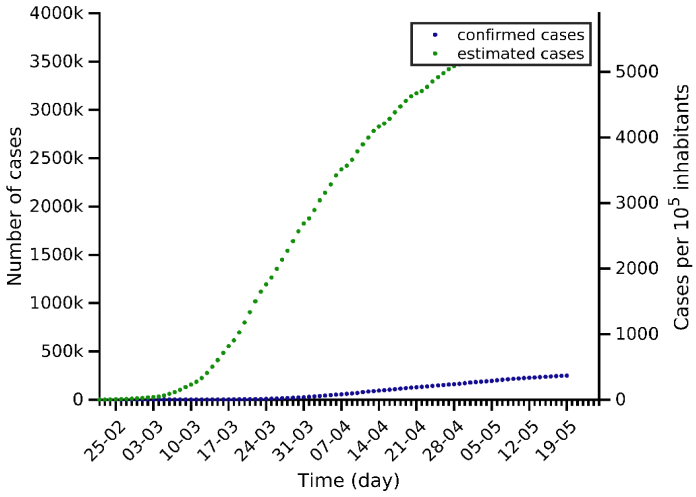


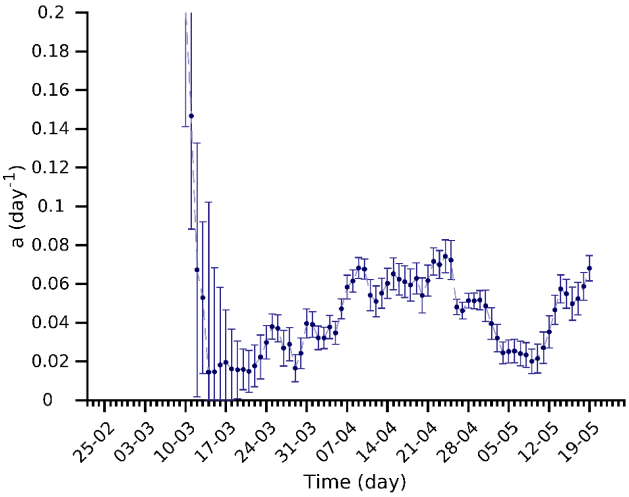


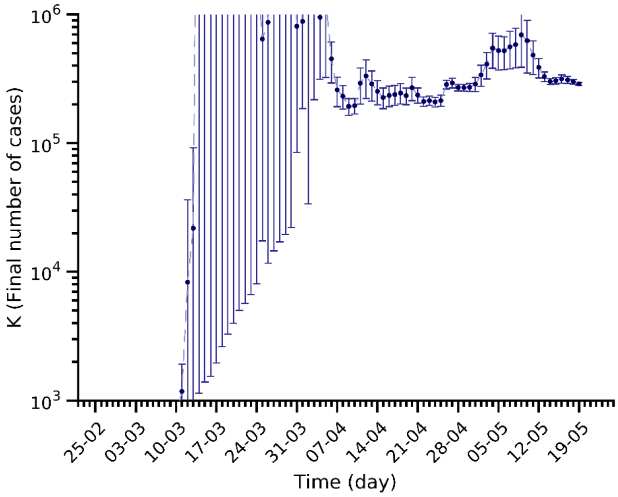


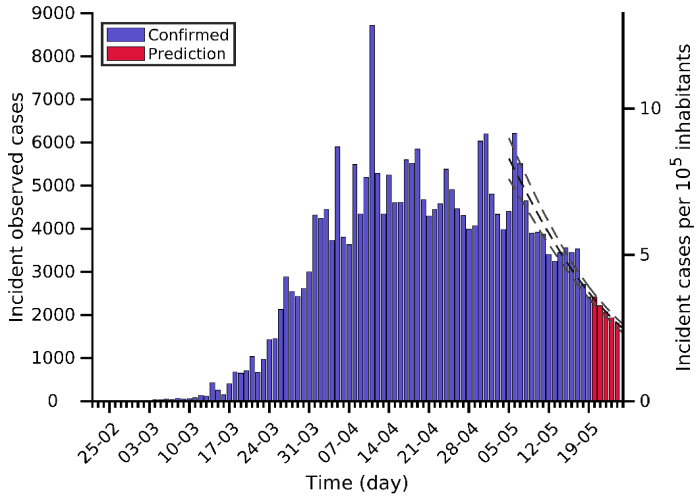
UK 19-05-2020. Population: 67.9M. Current cumulated incidence: $367/10^5$

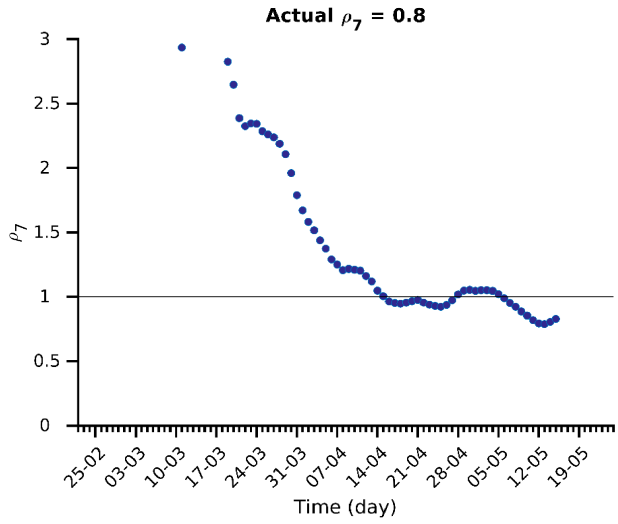


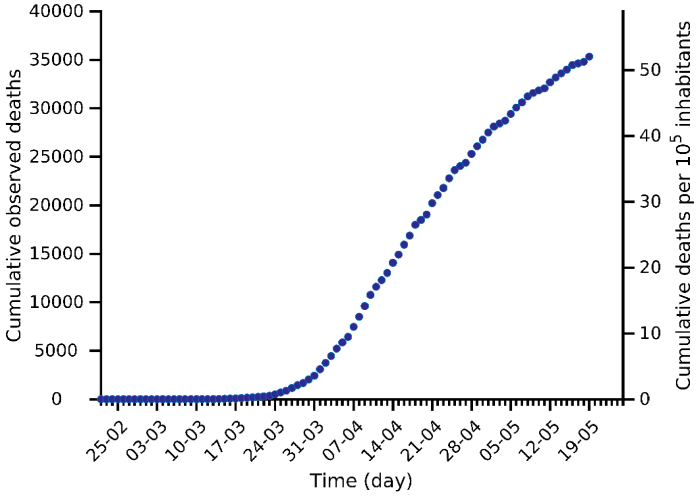


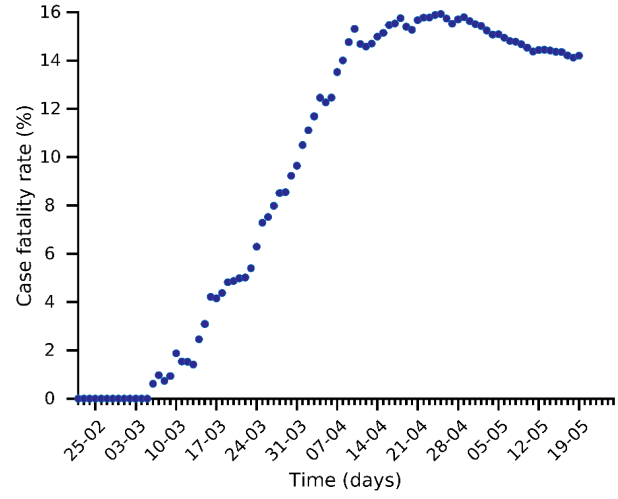




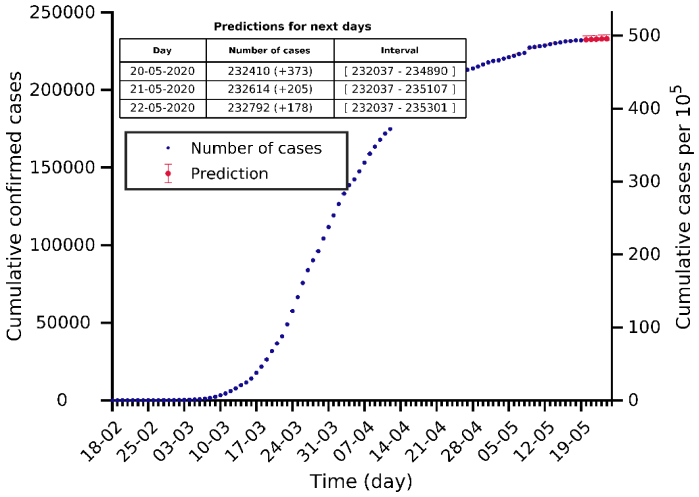


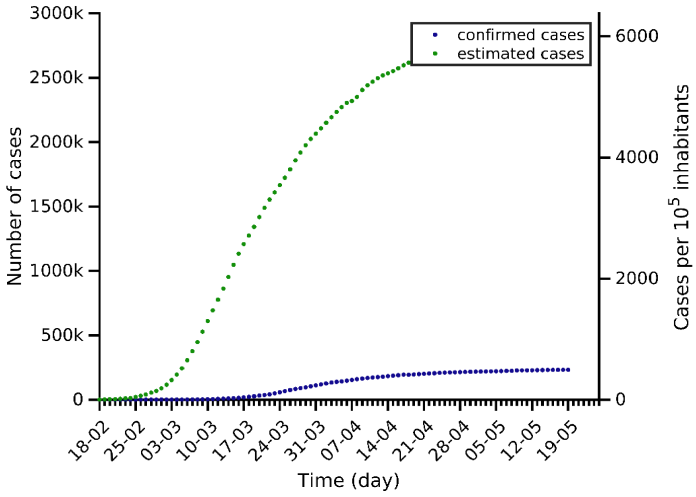


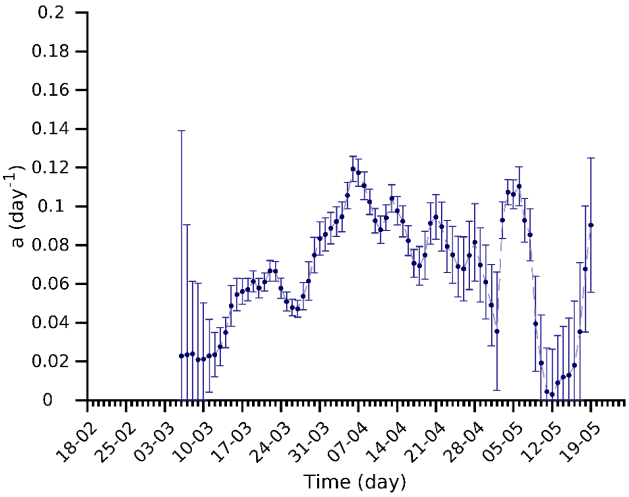


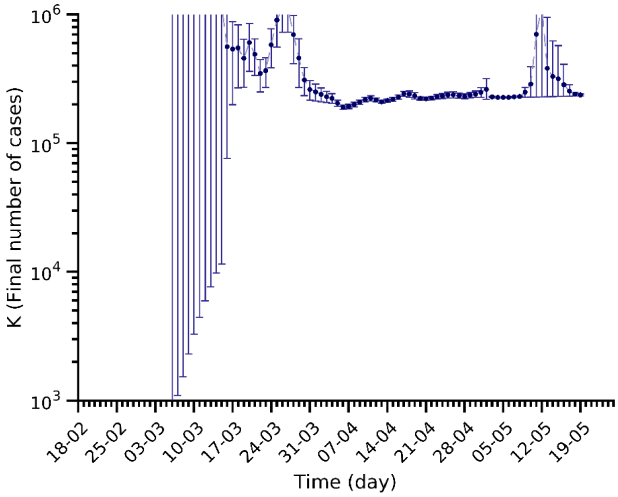


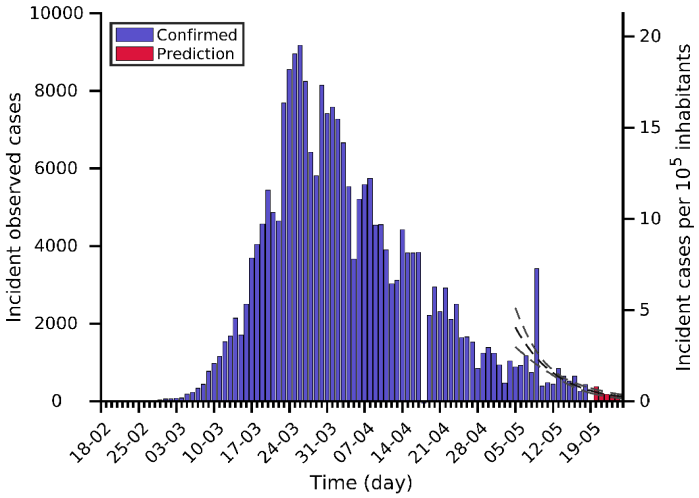
Spain 19-05-2020. Population: 47.0M. Current cumulated incidence: 493/10^5

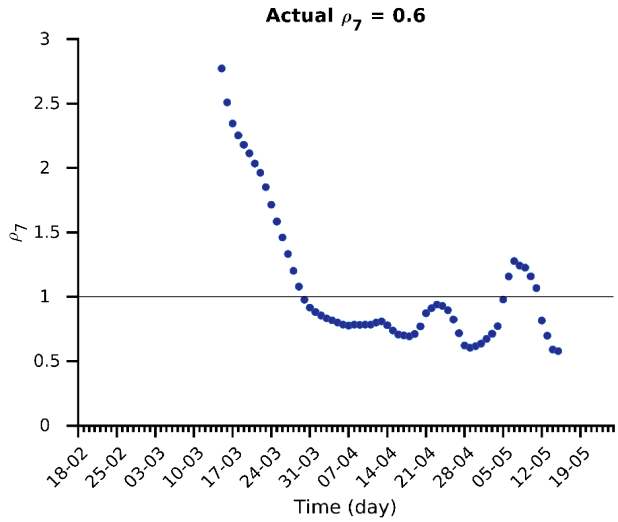


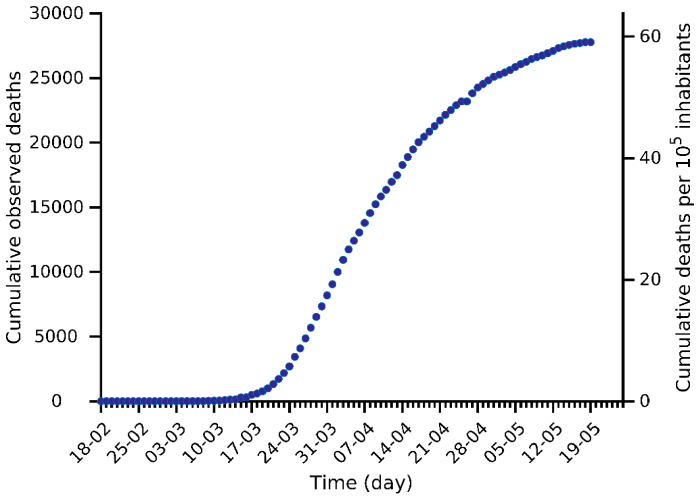


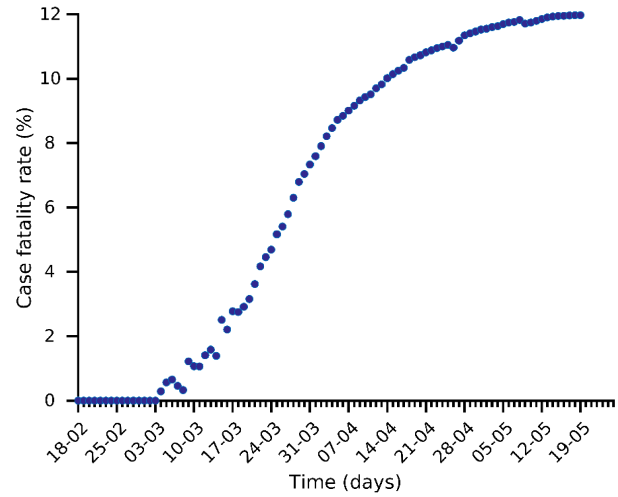








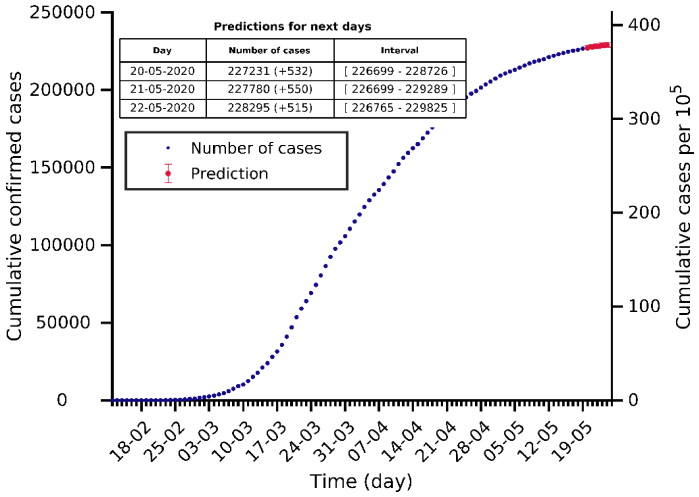


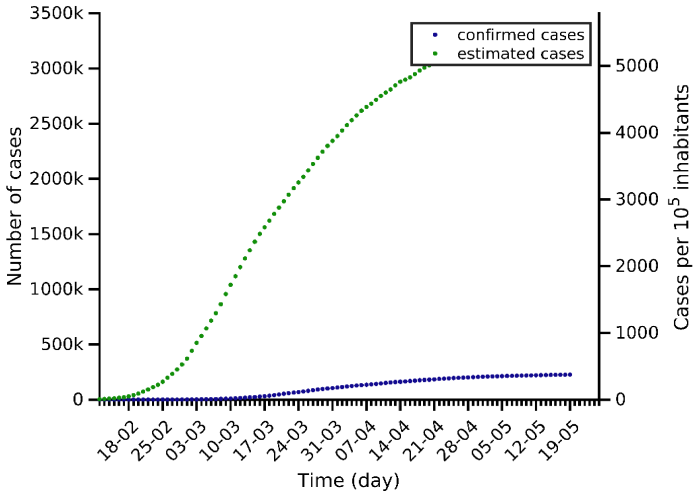


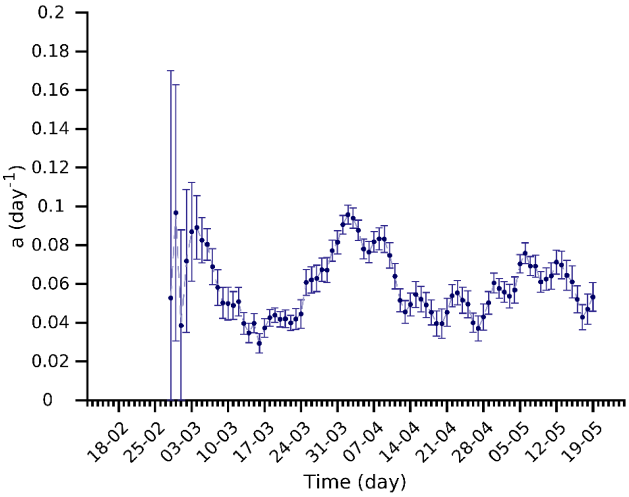
Italy 19-05-2020. Population: 60.5M. Current cumulated incidence: $375/10^5$

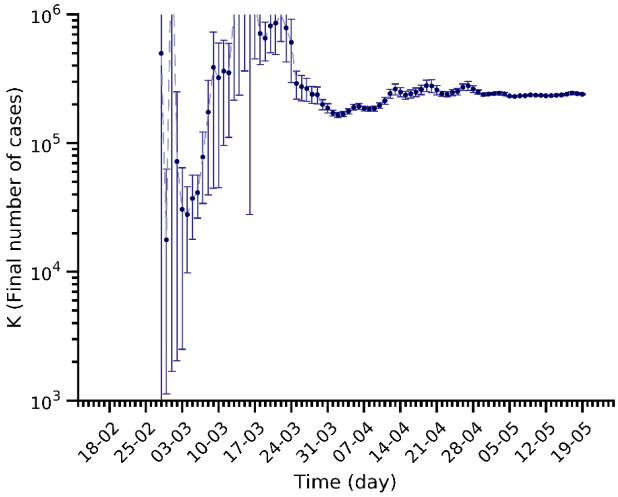
Predictions for next days

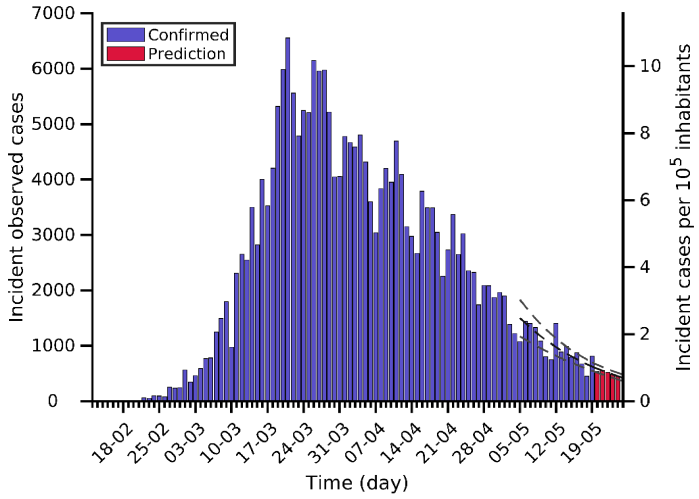
Day	Number of cases	Interval
20-05-2020	227231 (+532)	[226699 - 228726]
21-05-2020	227780 (+550)	[226699 - 229289]
22-05-2020	228295 (+515)	[226765 - 229825]

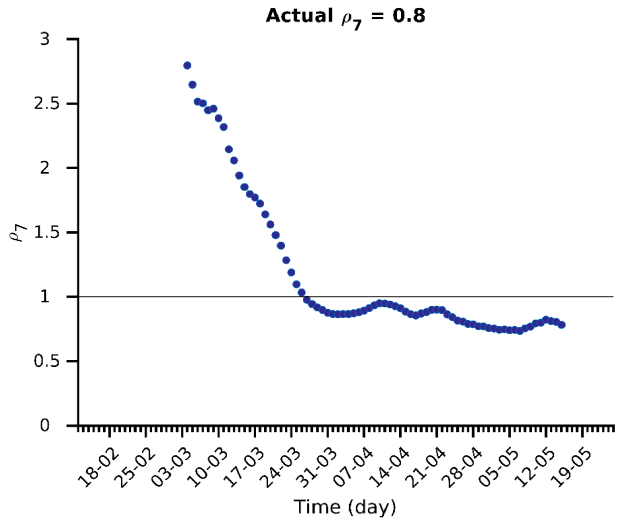


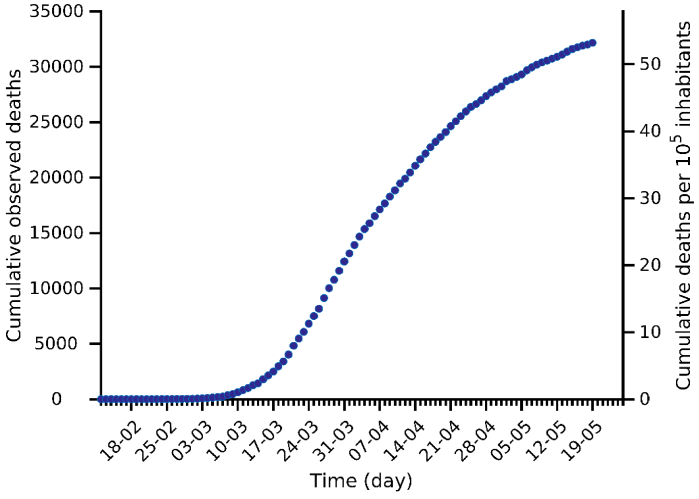


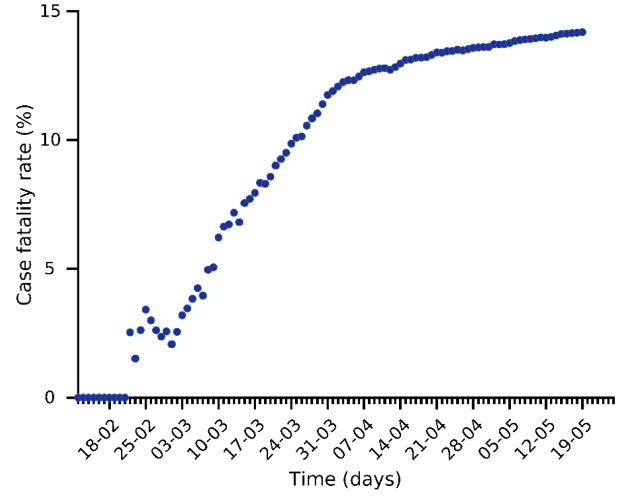




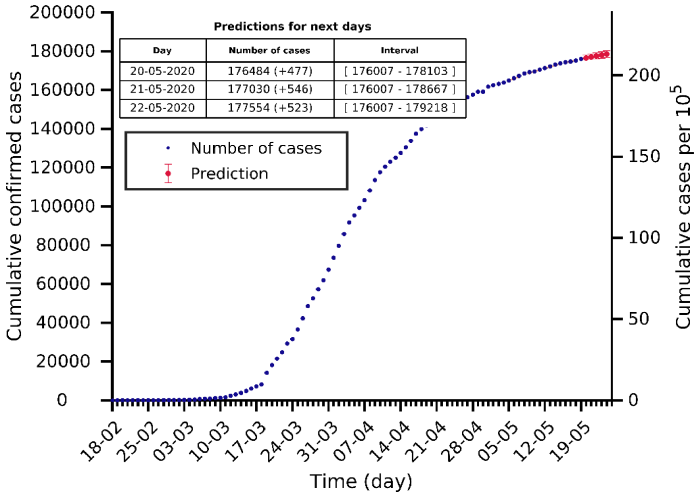


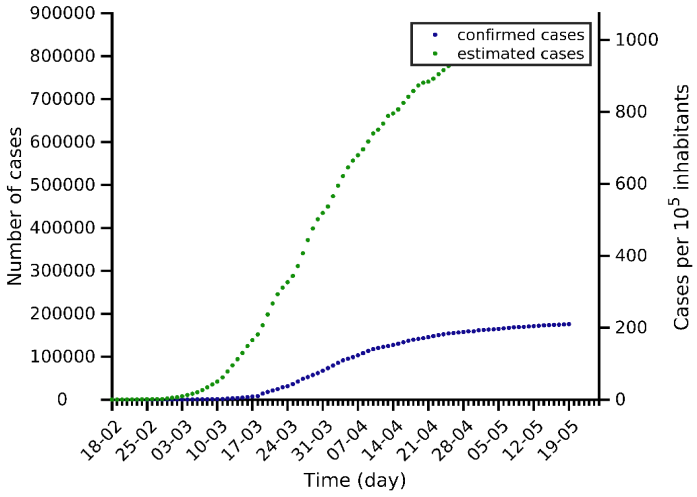


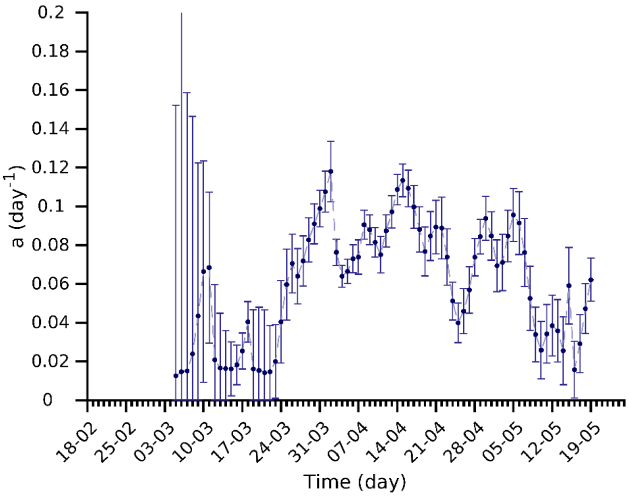


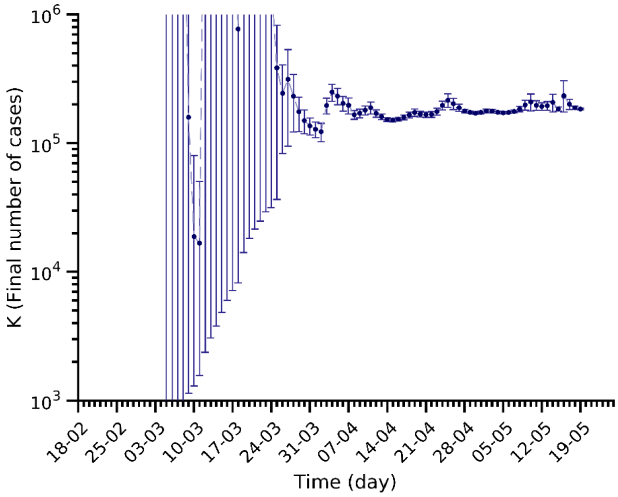


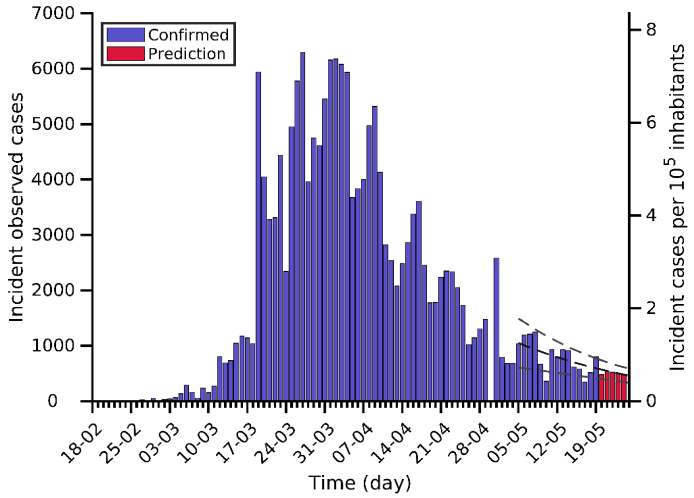
Germany 19-05-2020. Population: 83.8M. Current cumulated incidence: 210/10⁵

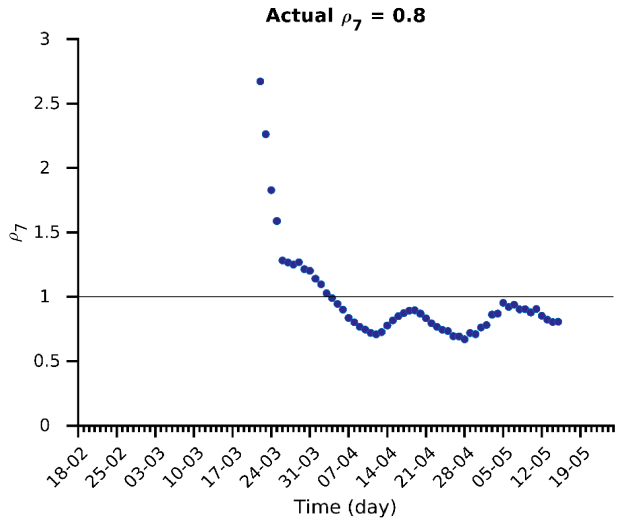


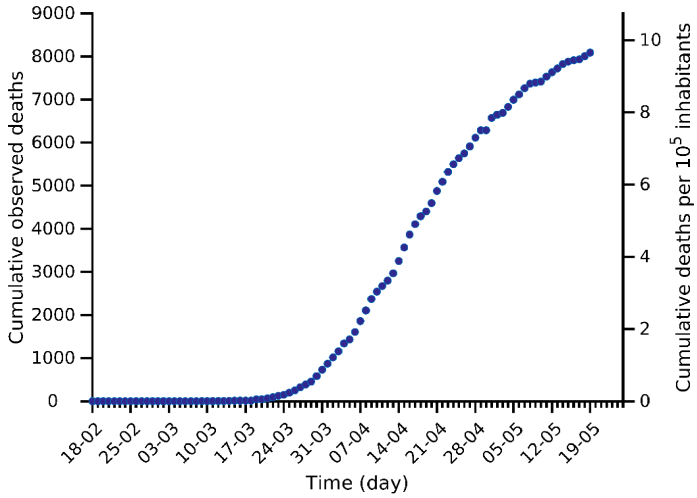


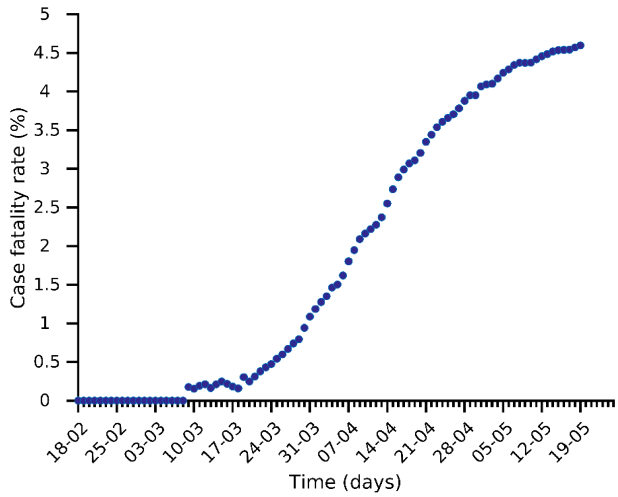




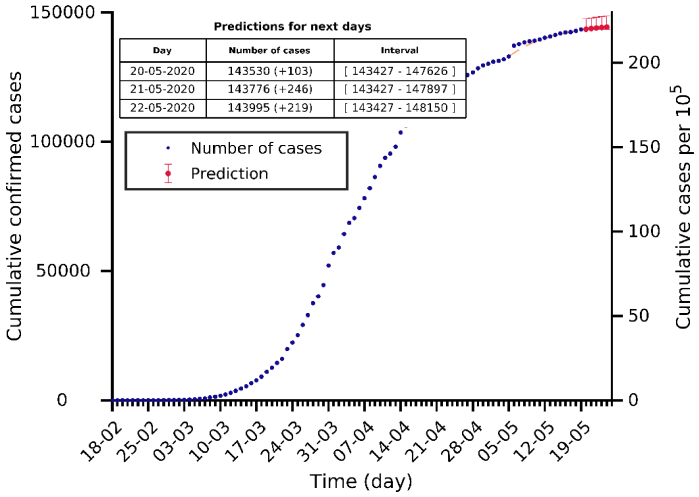


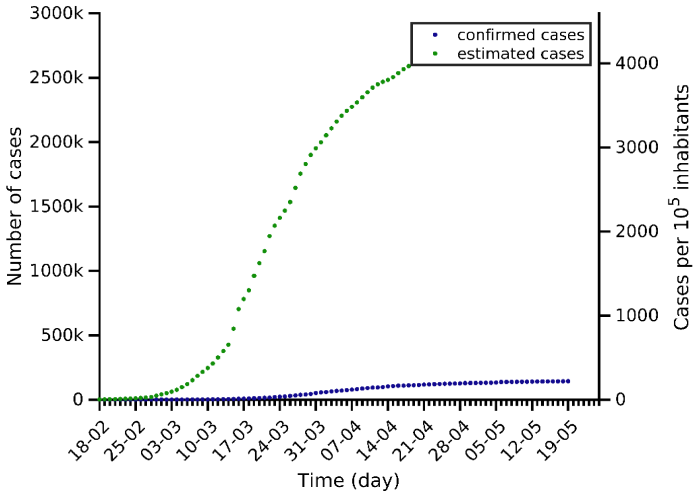


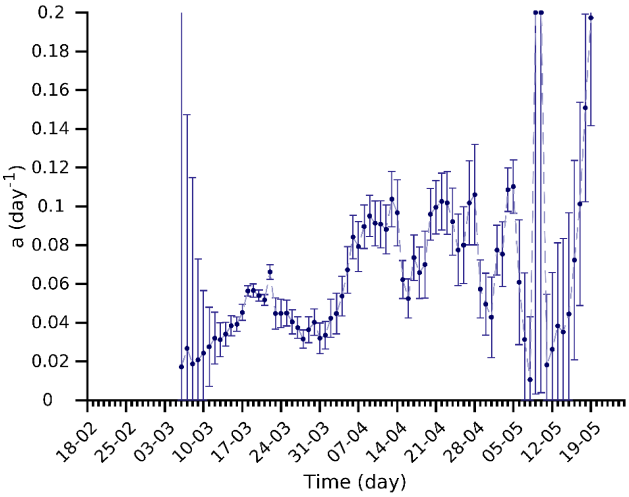


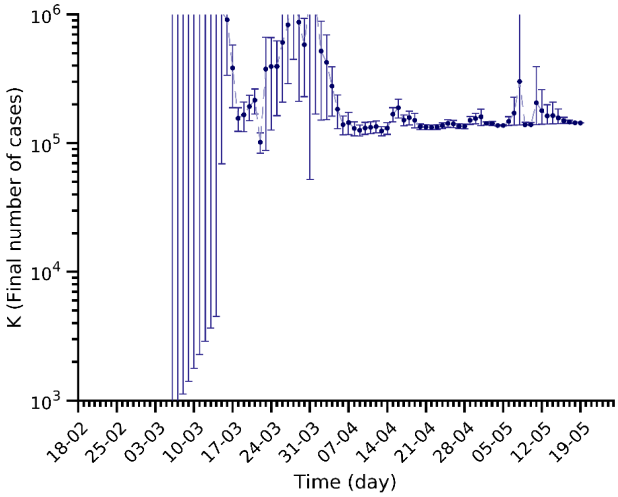


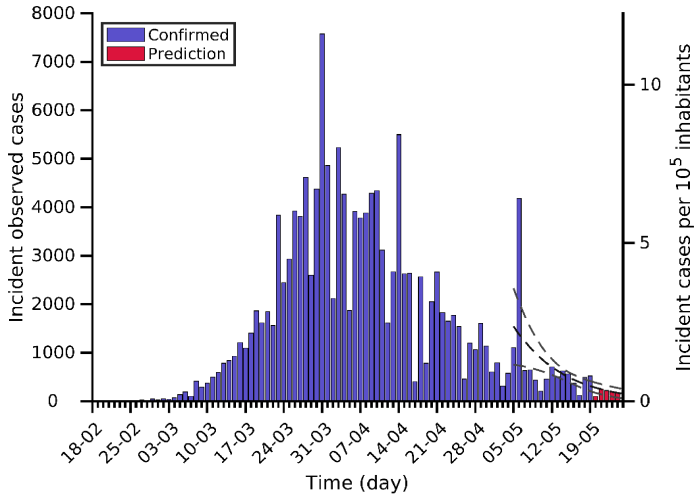
France 19-05-2020. Population: 65.3M. Current cumulated incidence: 220/10^5

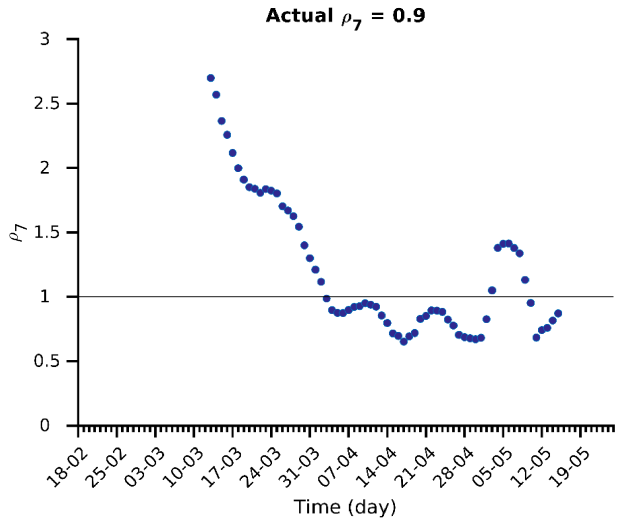


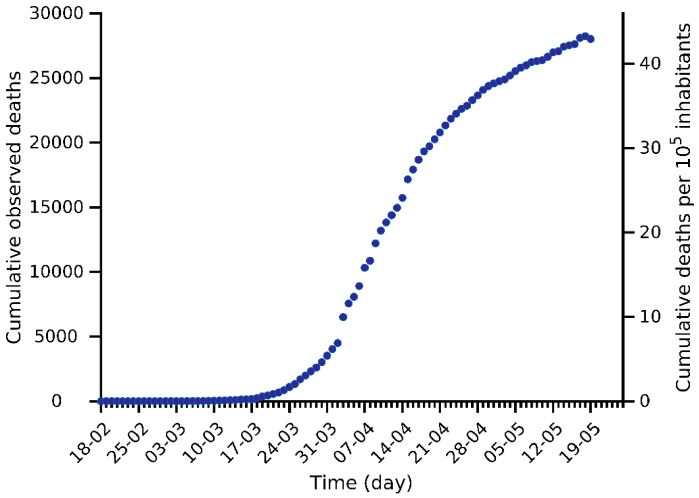


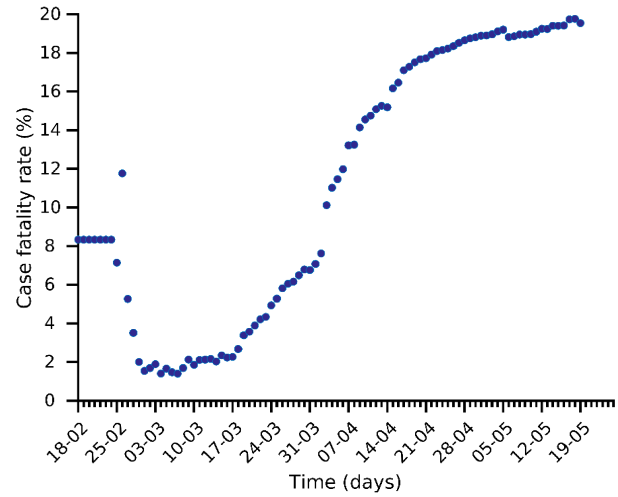




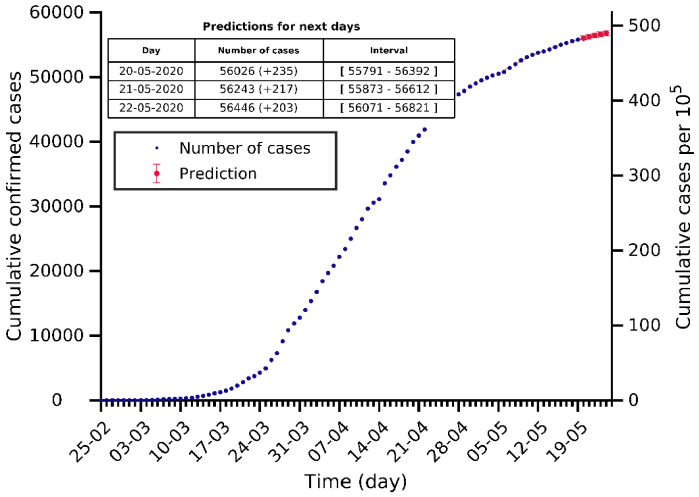


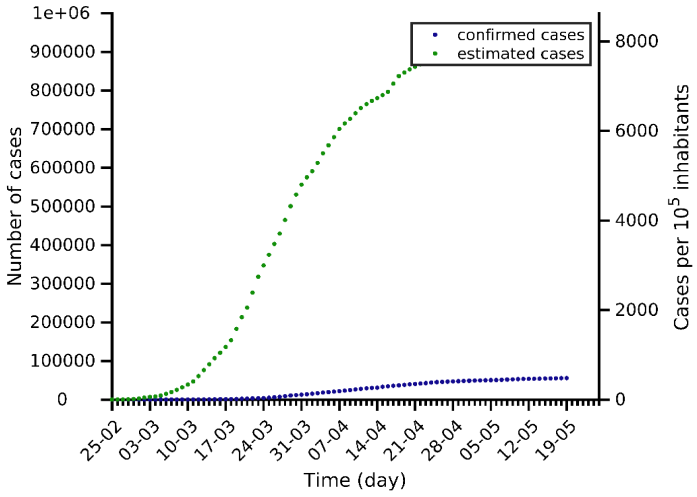


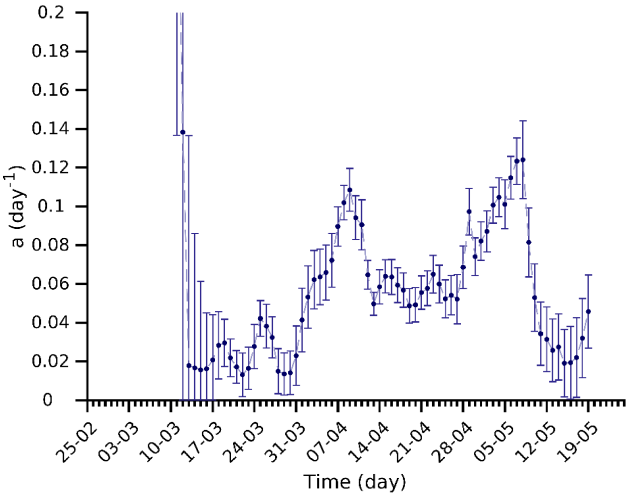


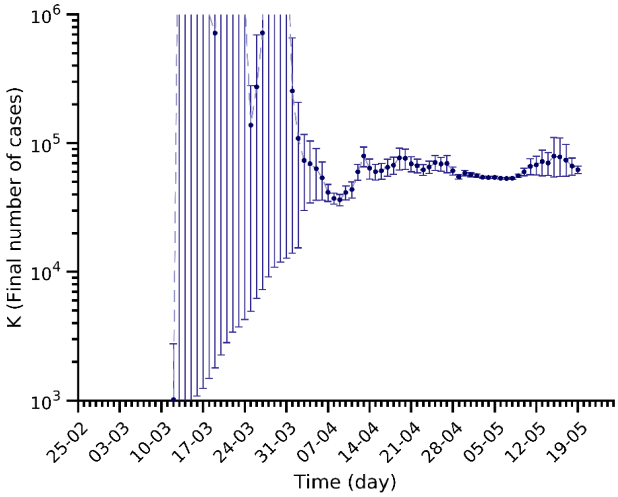


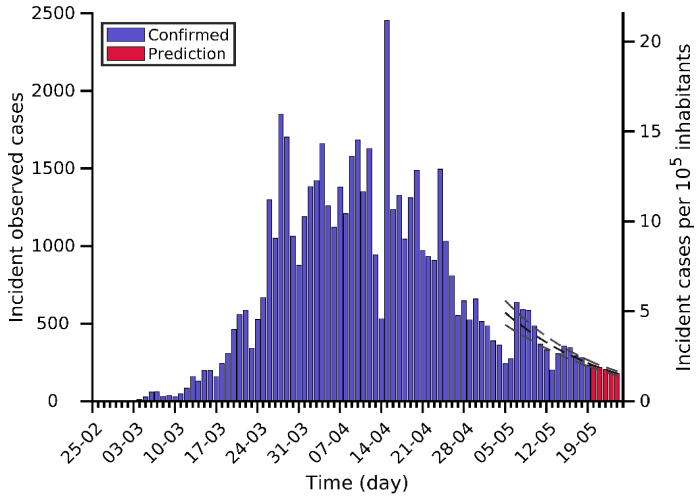
Belgium 19-05-2020. Population: 11.6M. Current cumulated incidence: 481/10^5

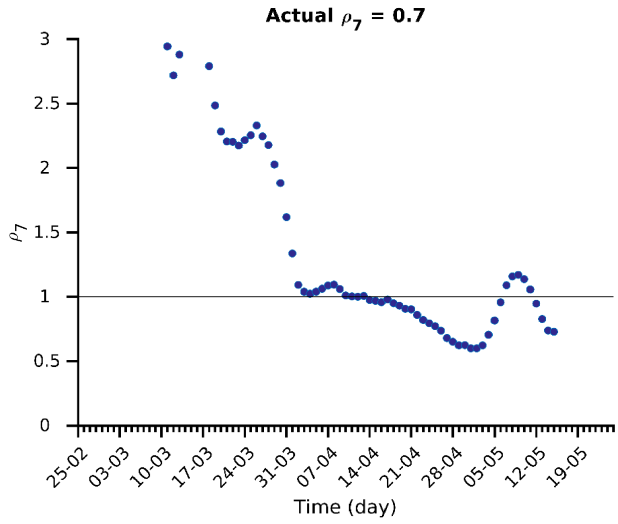


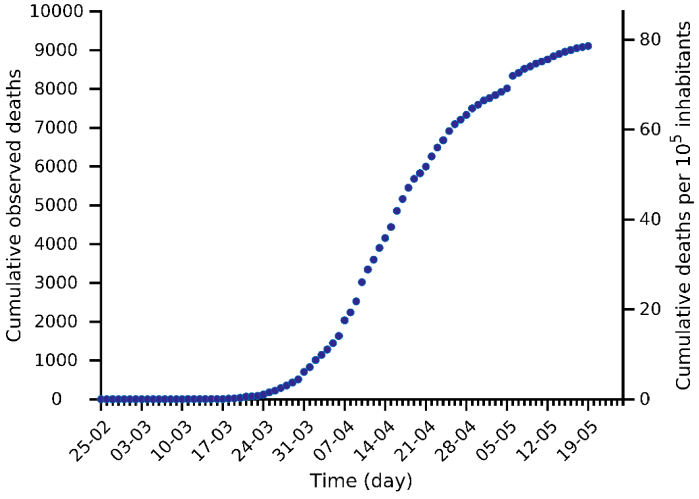


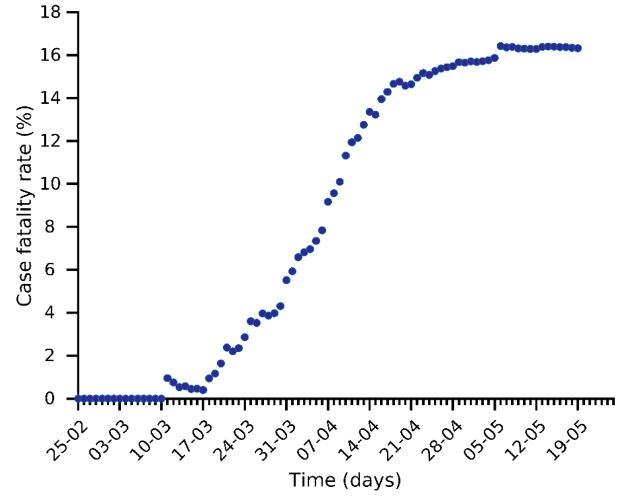




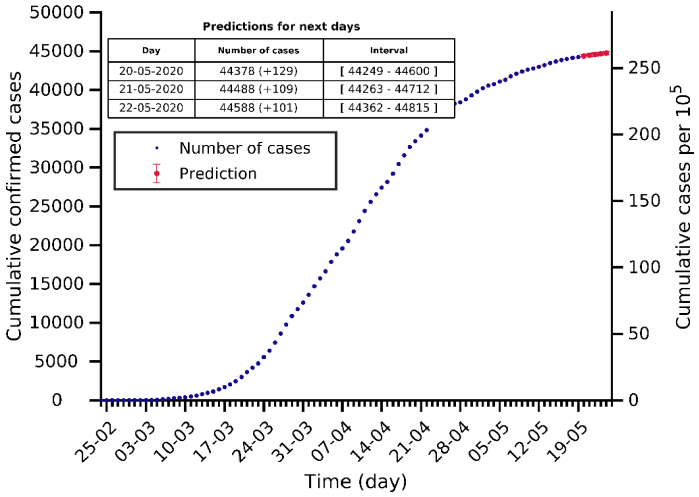


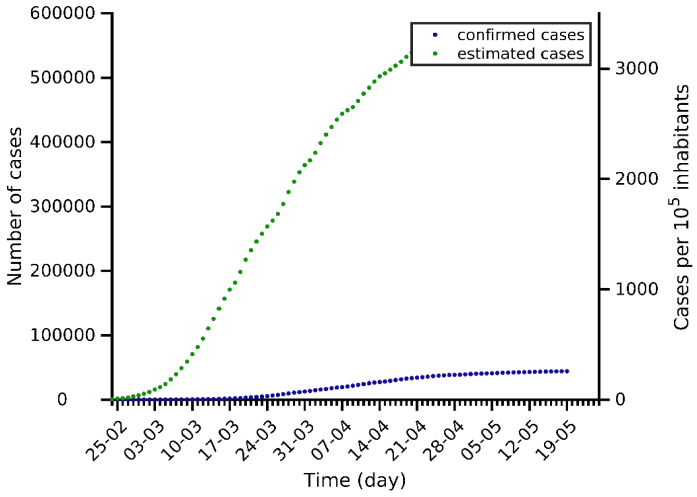


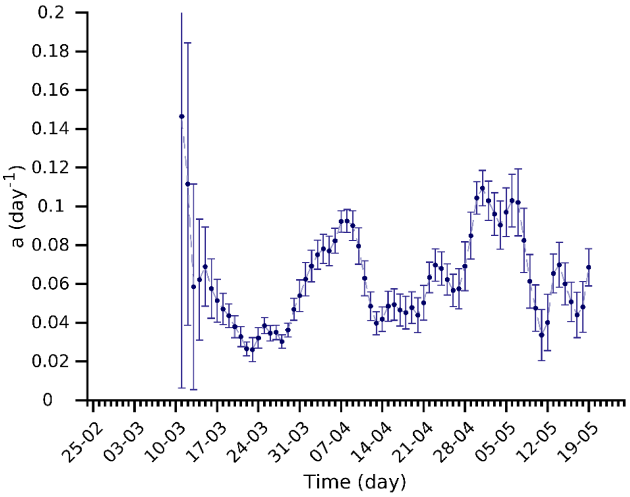


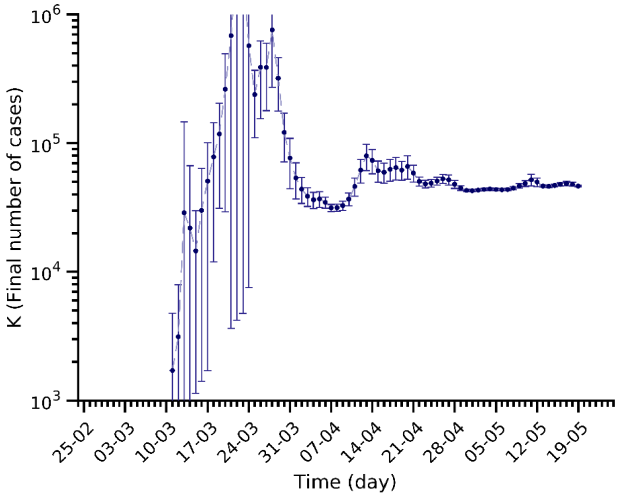


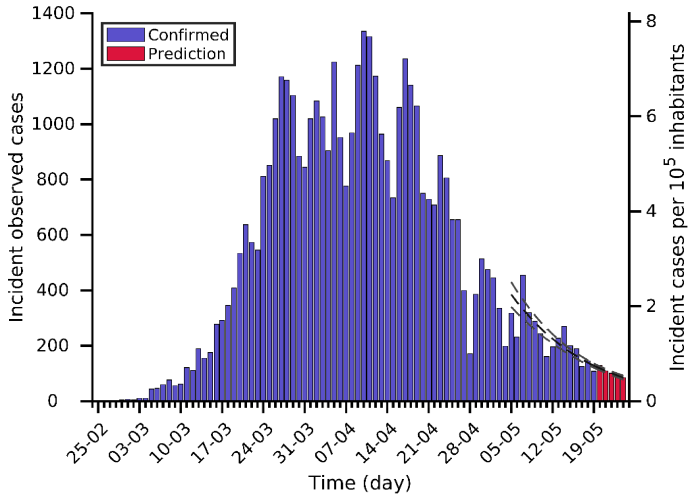
Netherlands 19-05-2020. Population: 17.1M. Current cumulated incidence: 258/10^5

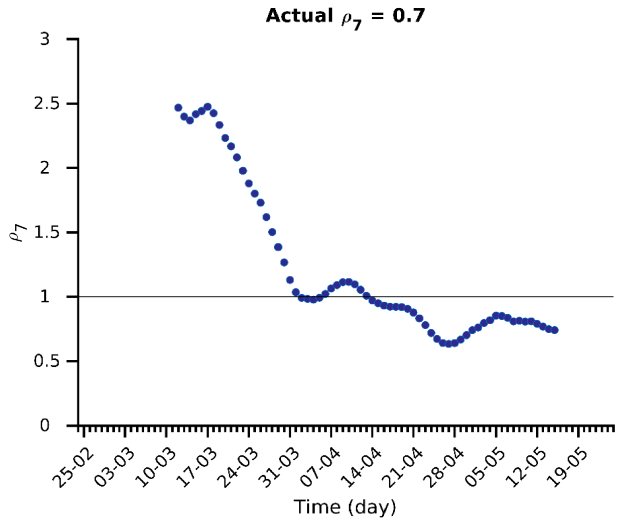


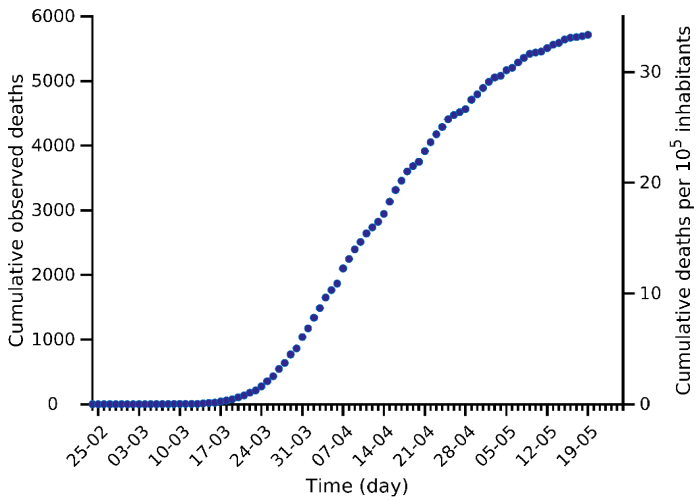


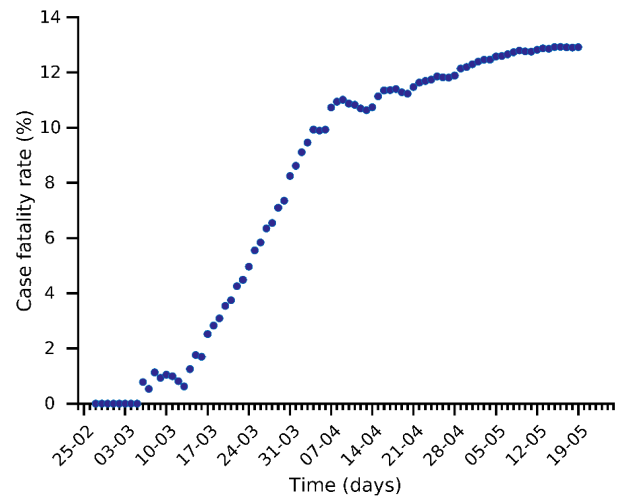




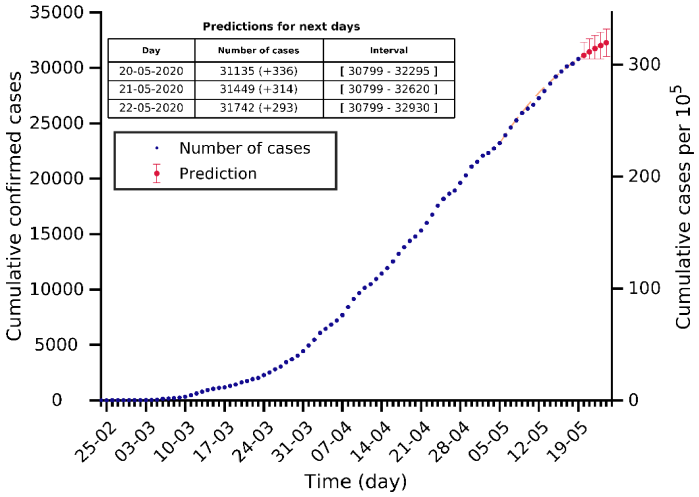


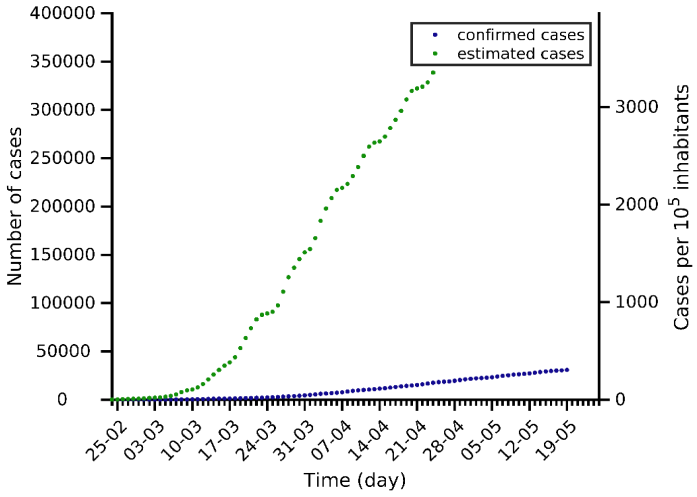


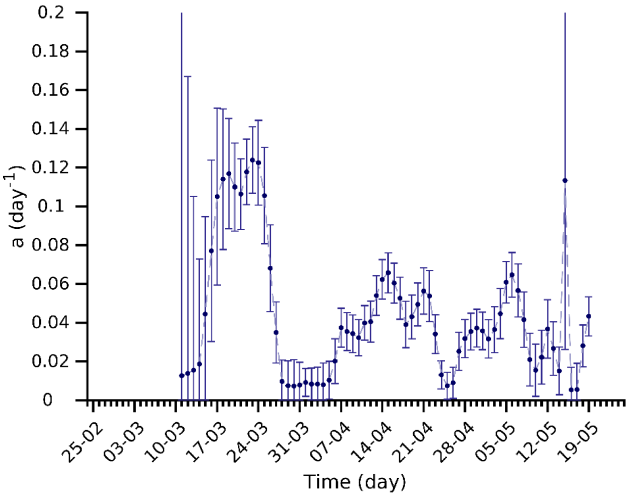


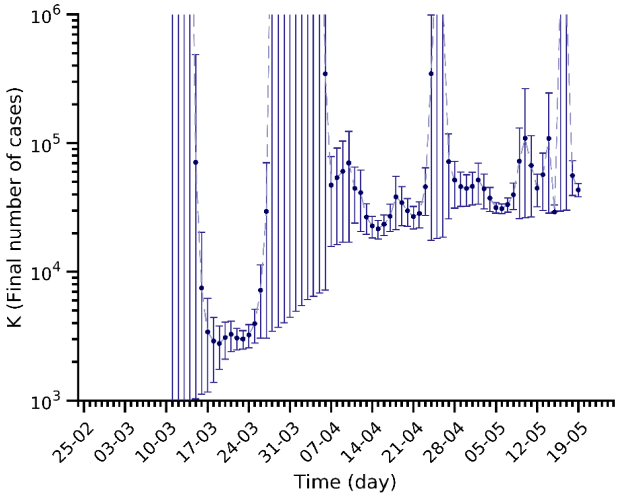


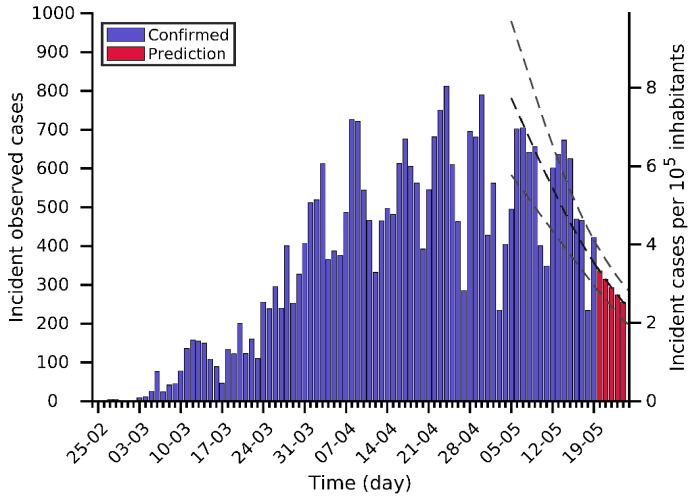
Sweden 19-05-2020. Population: 10.1M. Current cumulated incidence: $305/10^5$

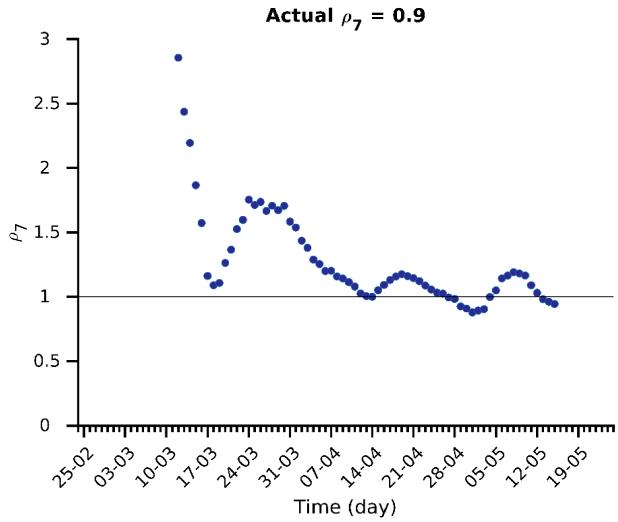


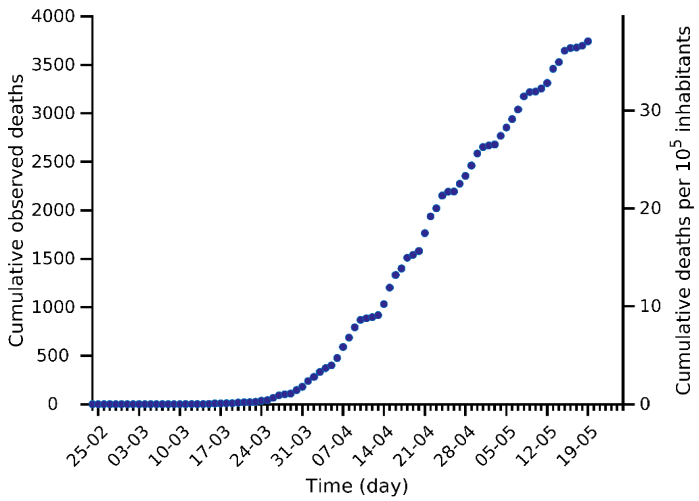


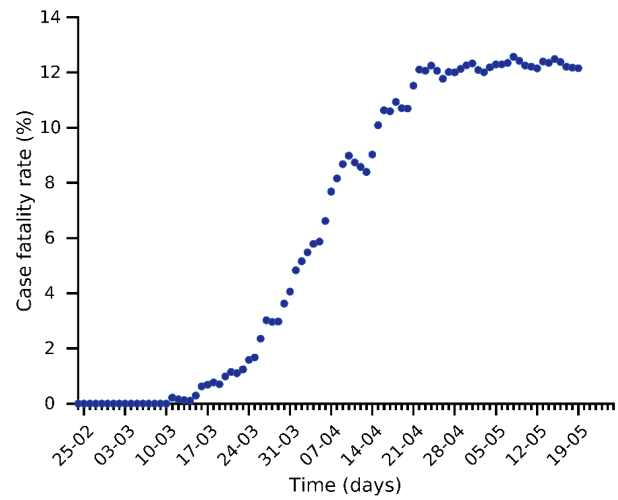




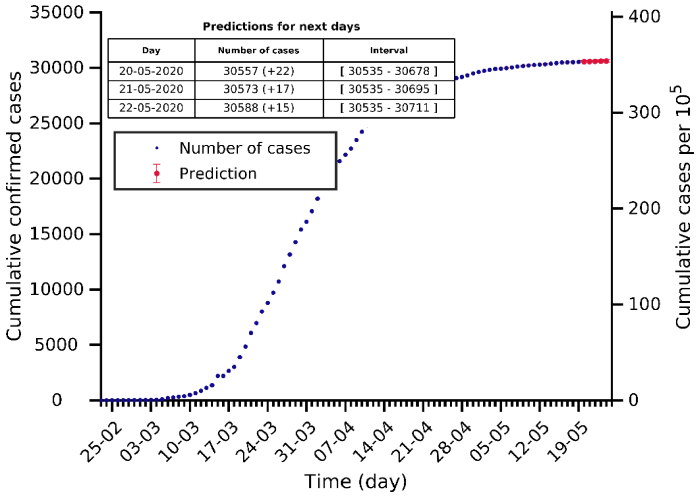


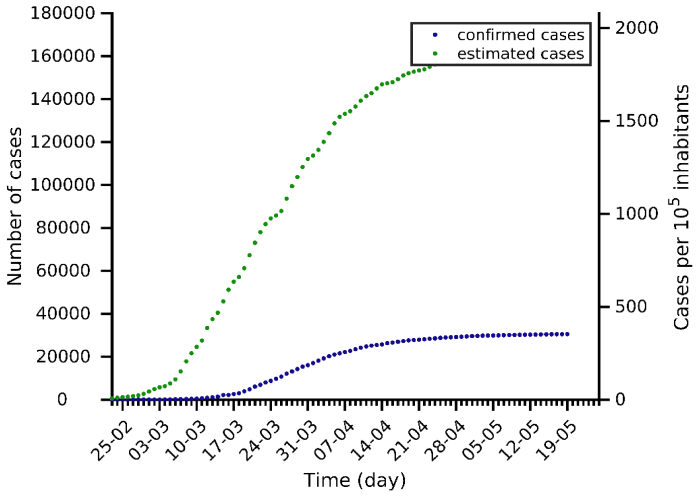


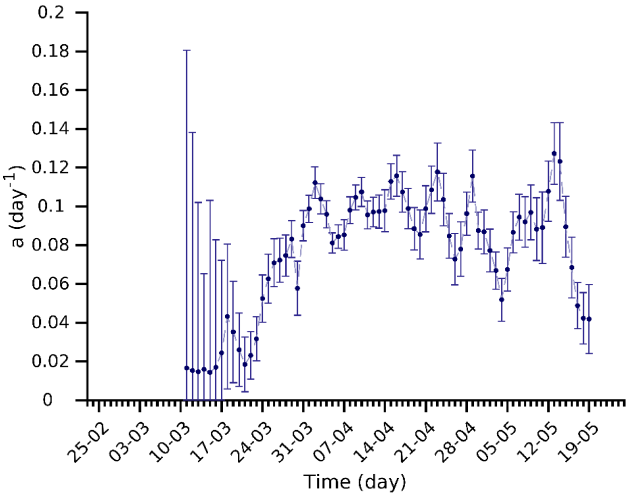


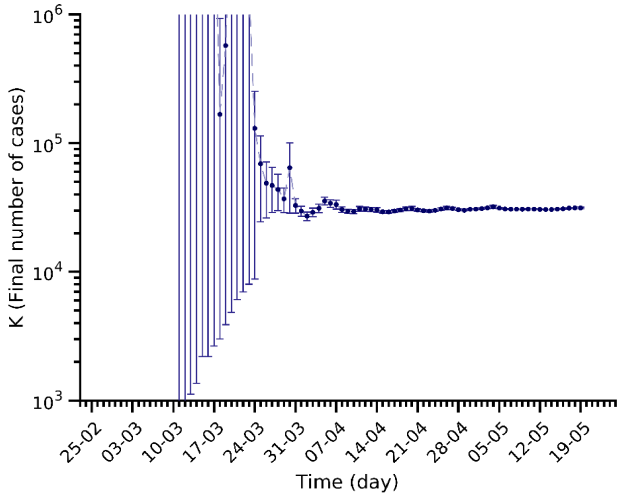


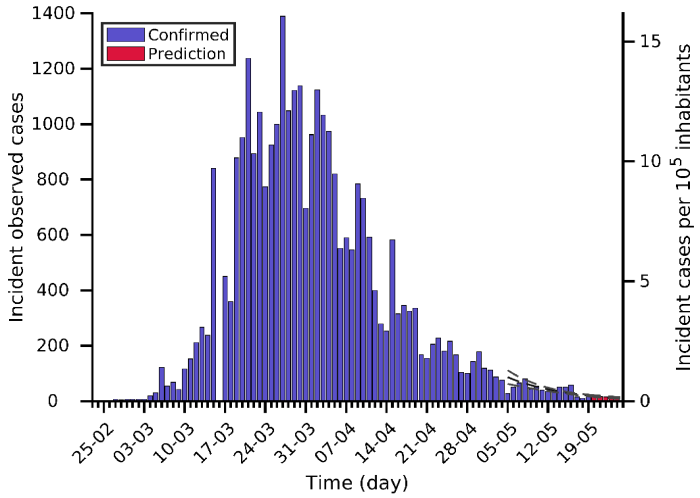
Switzerland 19-05-2020. Population: 8.7M. Current cumulated incidence: 353/10⁵

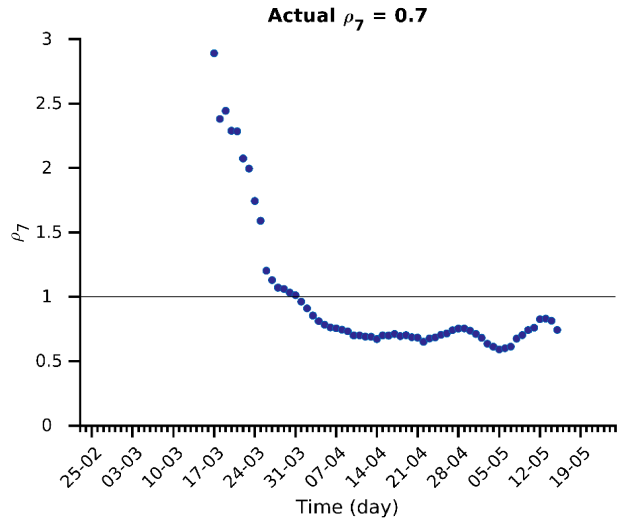


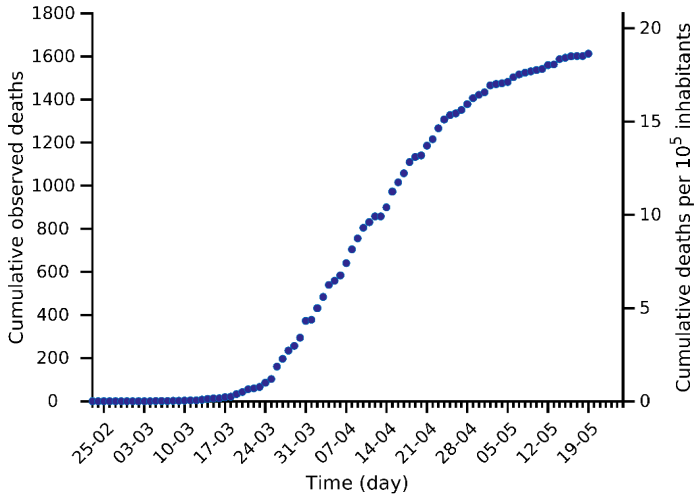


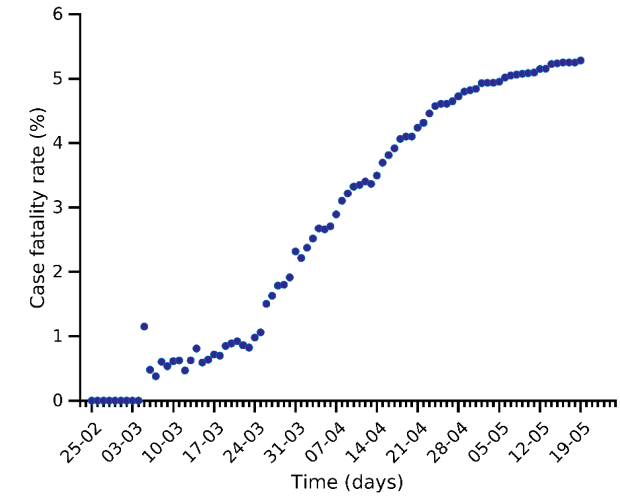




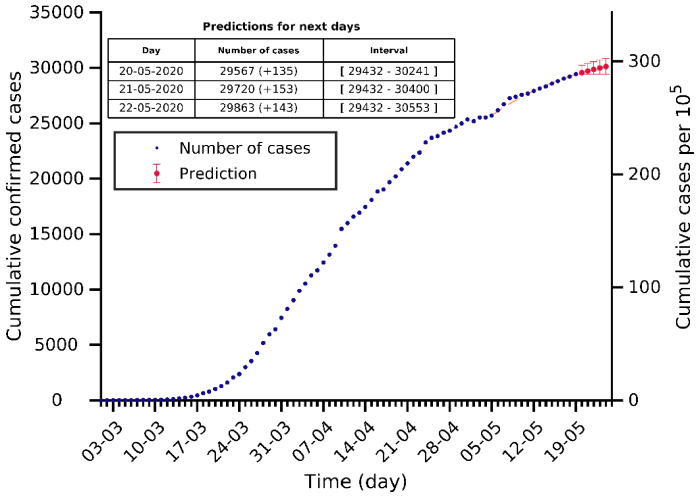


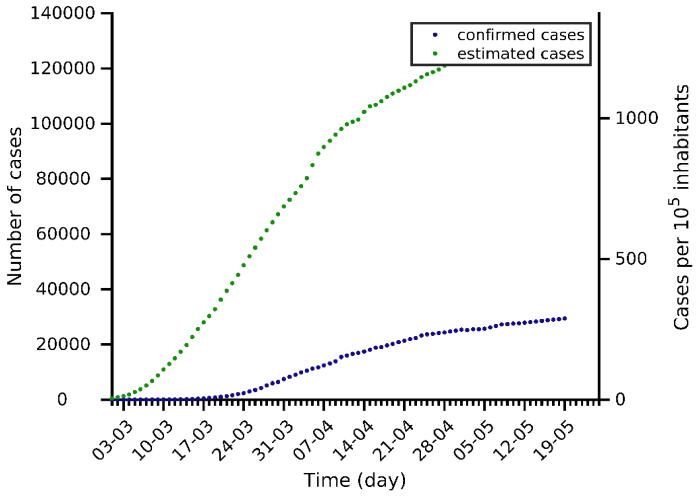


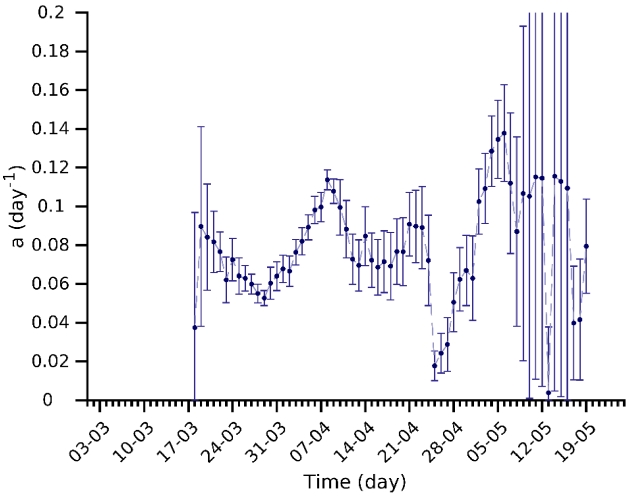


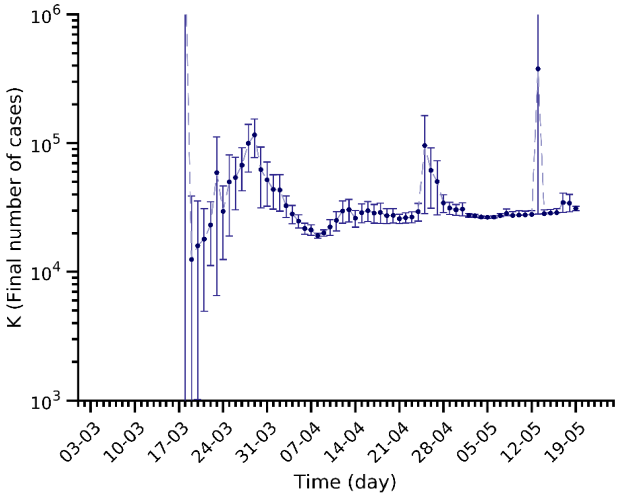


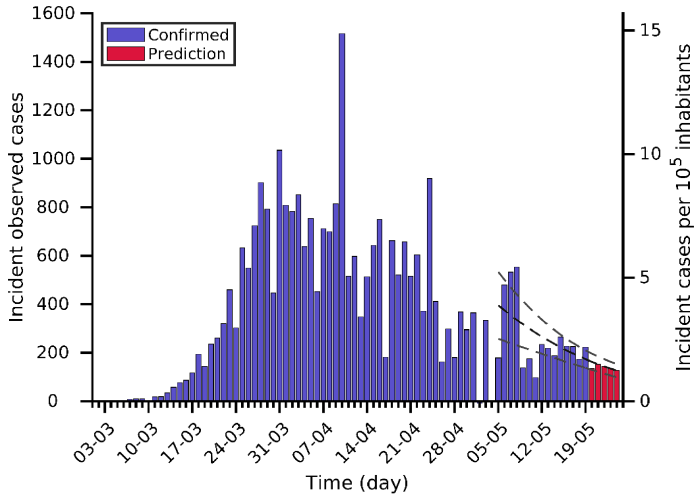
Portugal 19-05-2020. Population: 10.2M. Current cumulated incidence: 289/10⁵

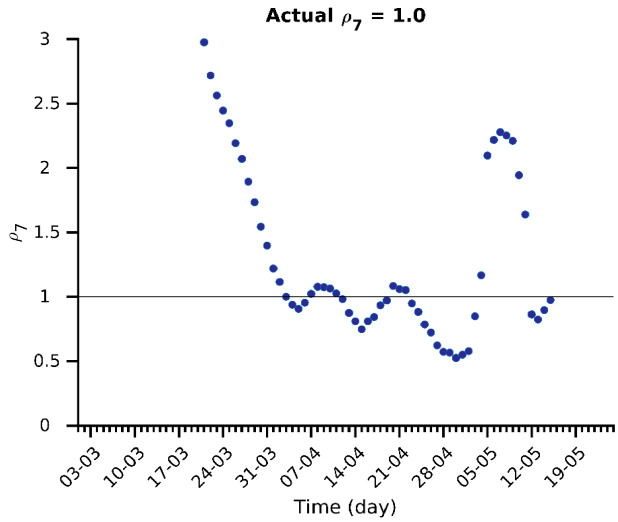


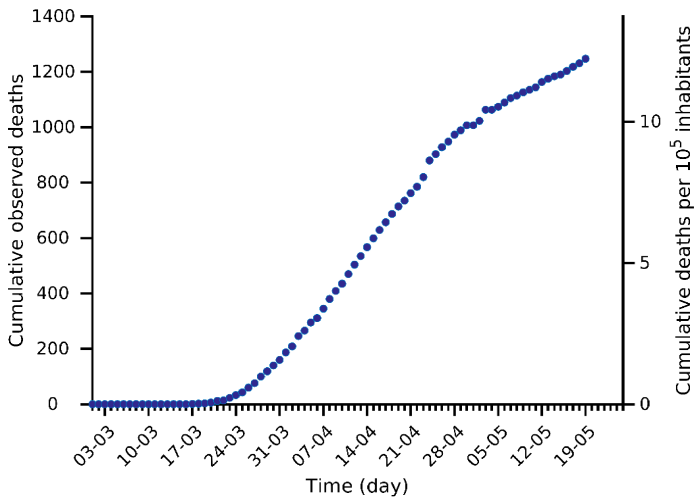


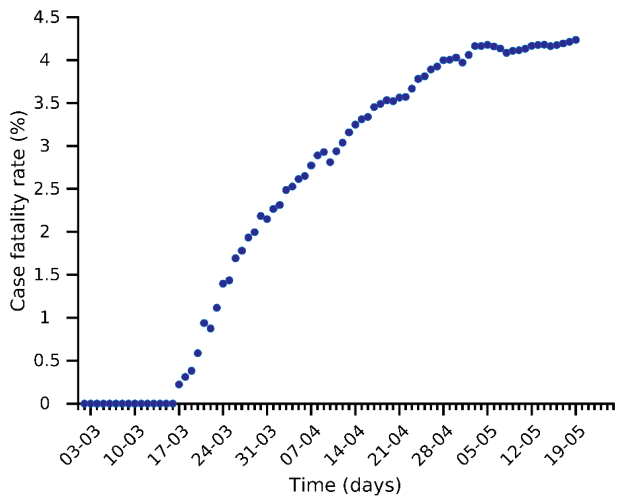




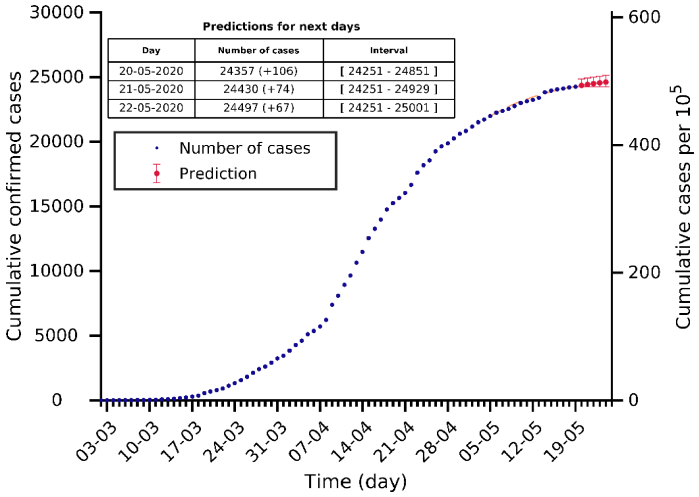


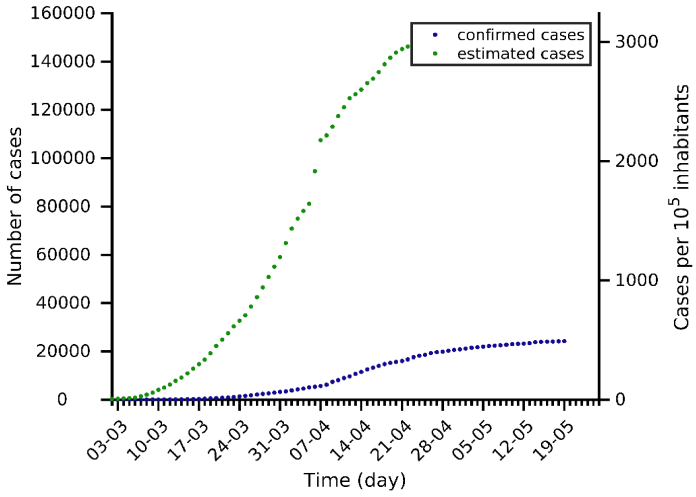


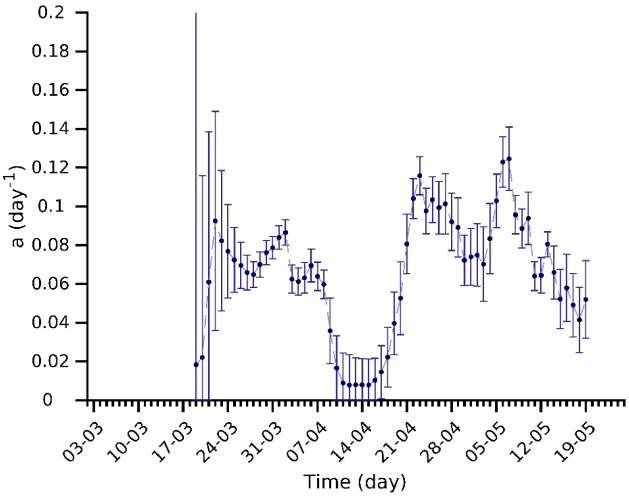


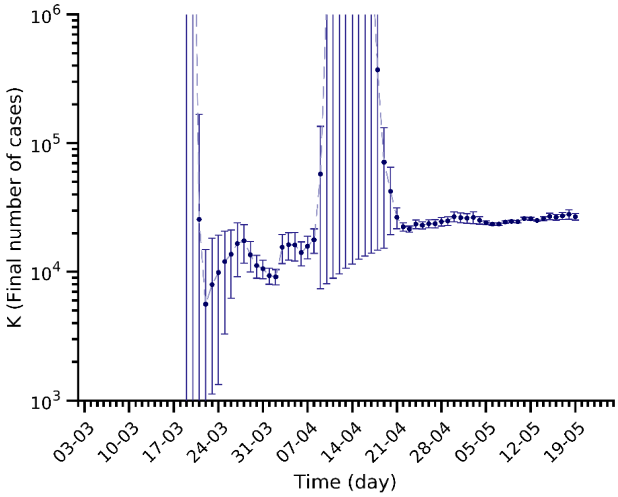


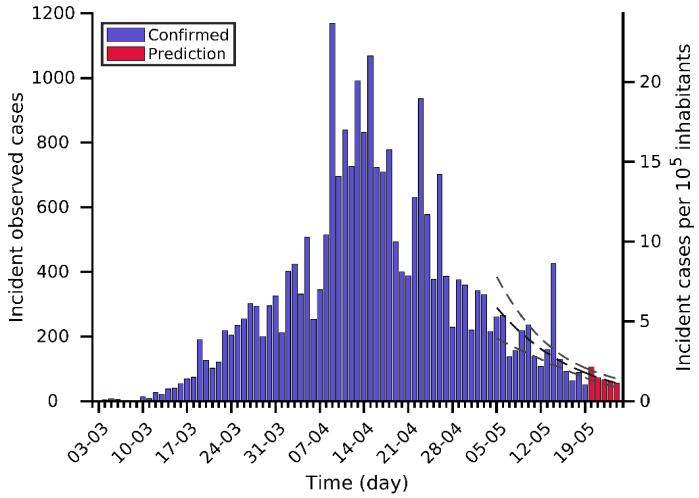
Ireland 19-05-2020. Population: 4.9M. Current cumulated incidence: 491/10⁵

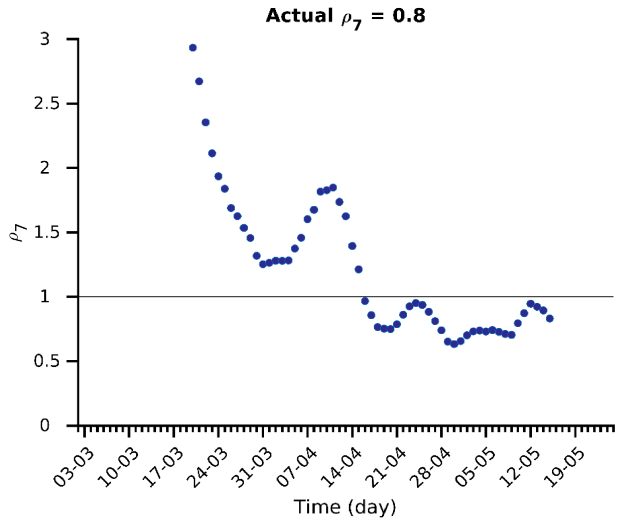


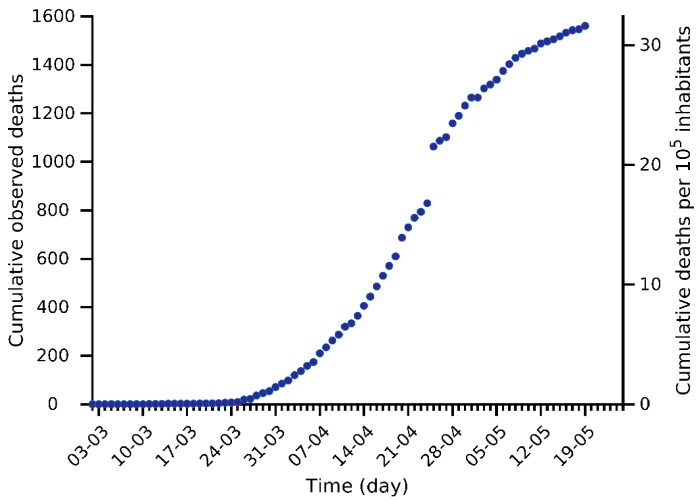


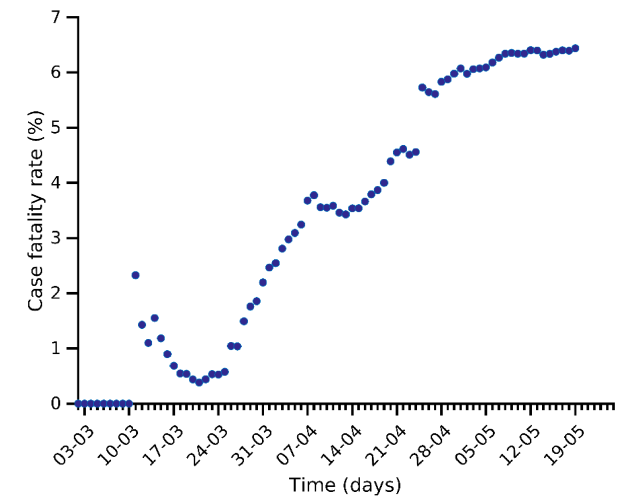




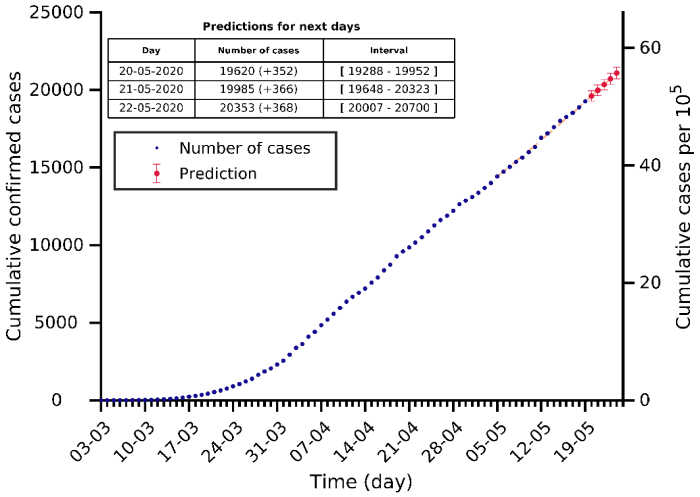


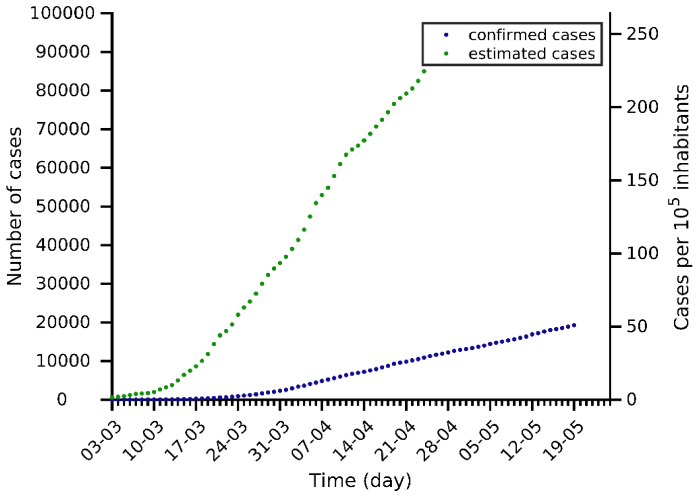


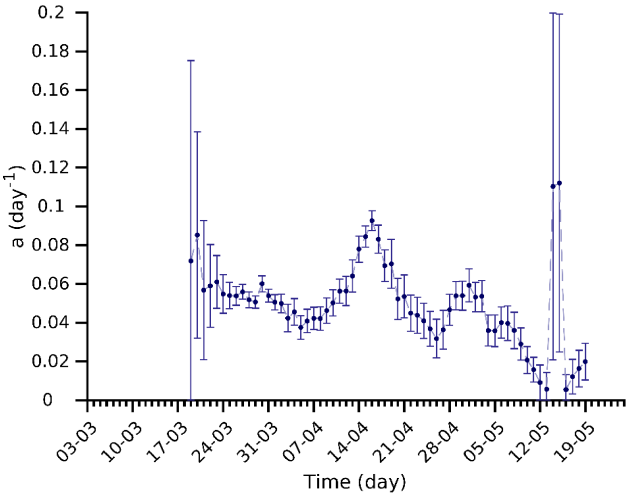


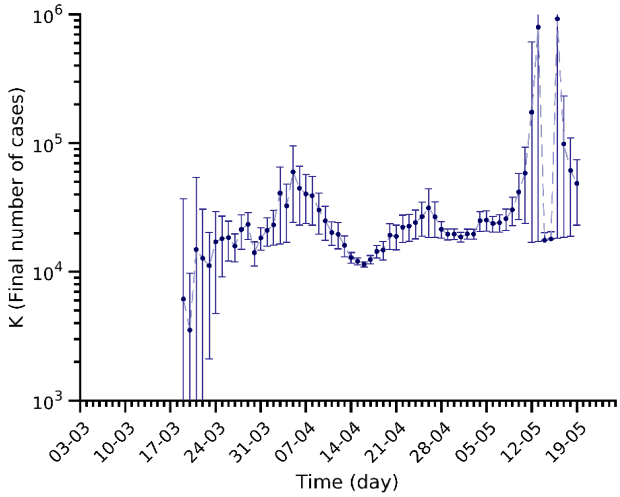


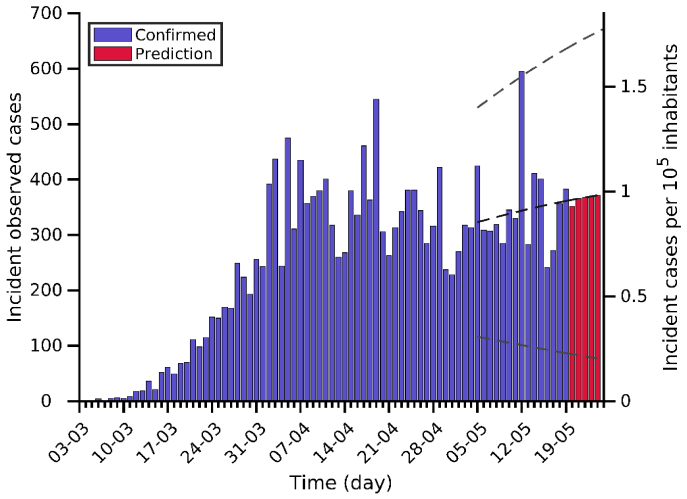
Poland 19-05-2020. Population: 37.8M. Current cumulated incidence: 51/10^5

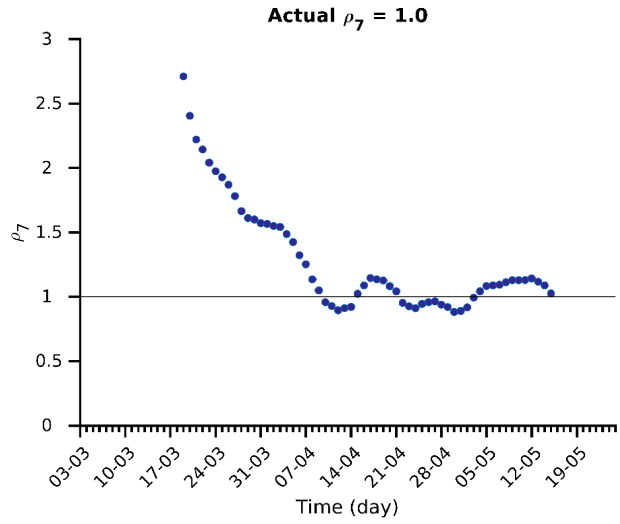


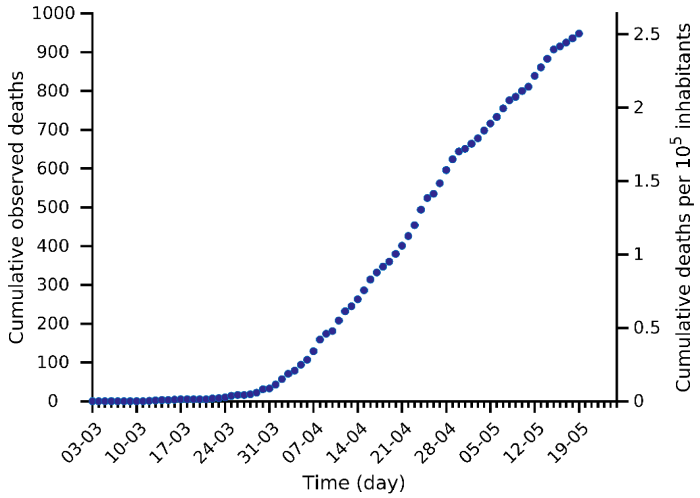


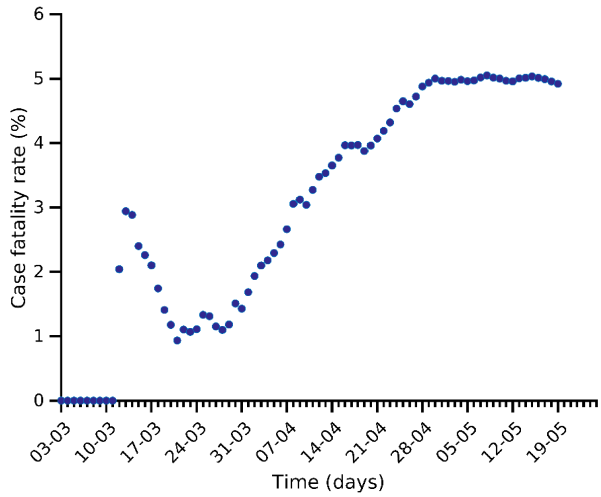




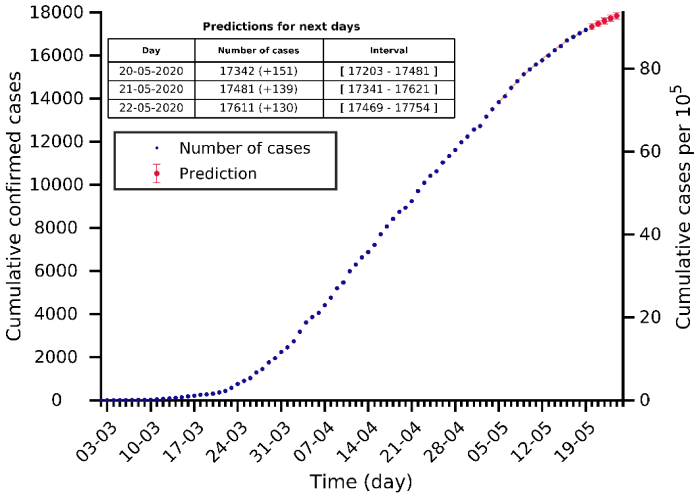


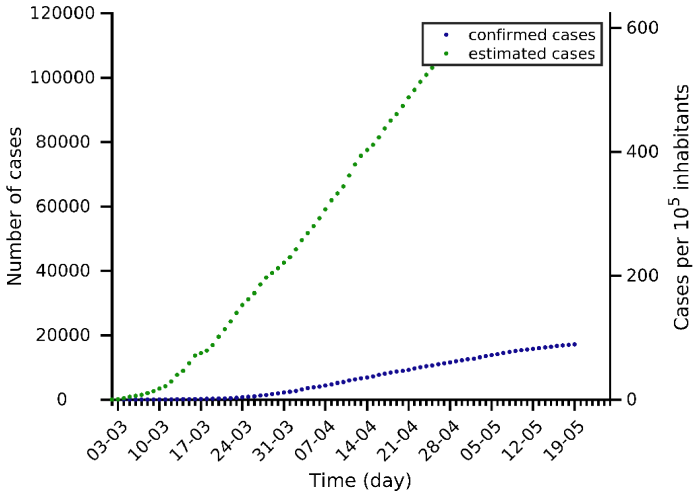


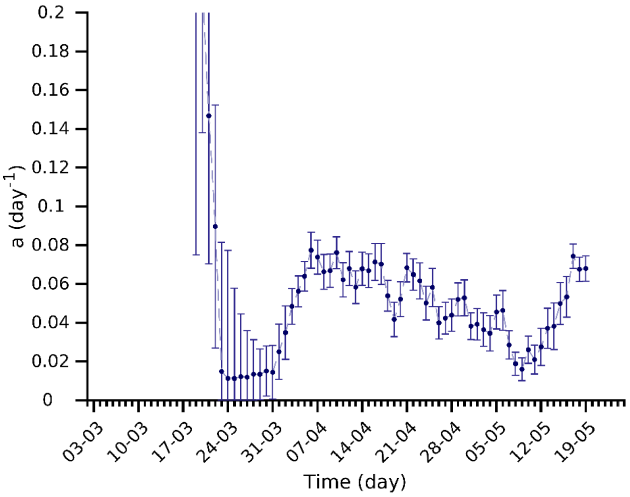


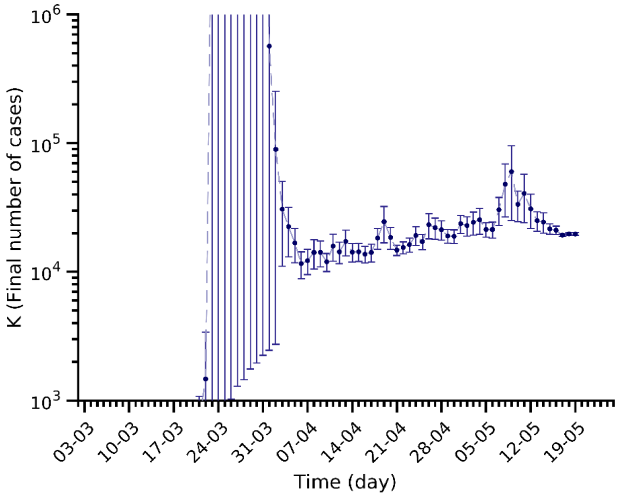


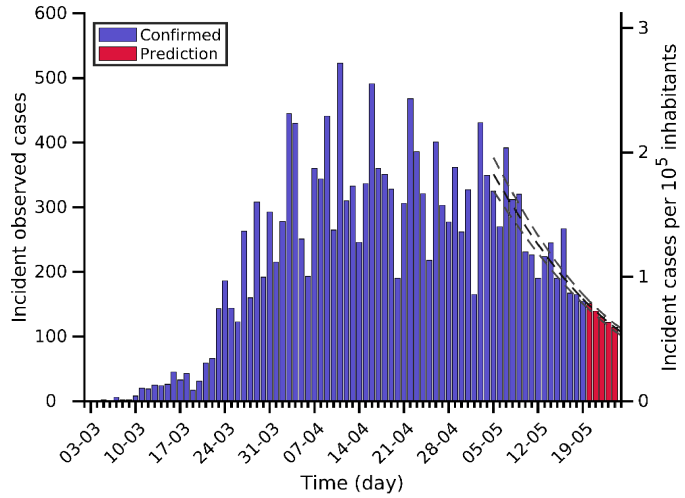
Romania 19-05-2020. Population: 19.2M. Current cumulated incidence: $89/10^5$

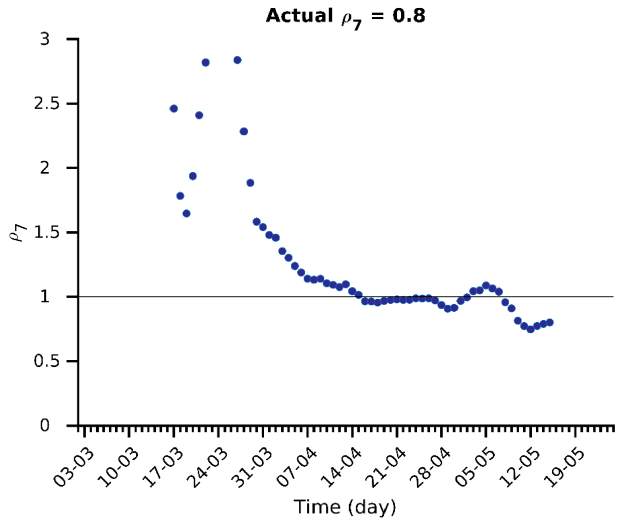


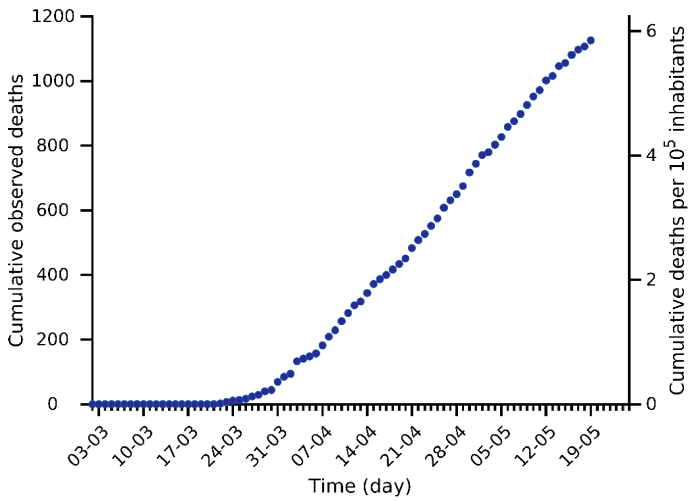


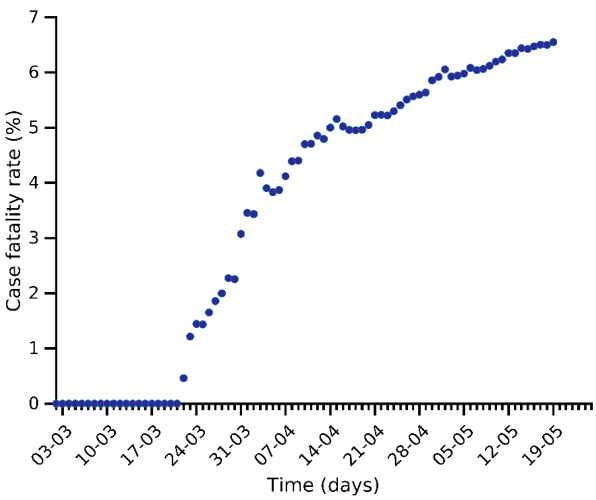




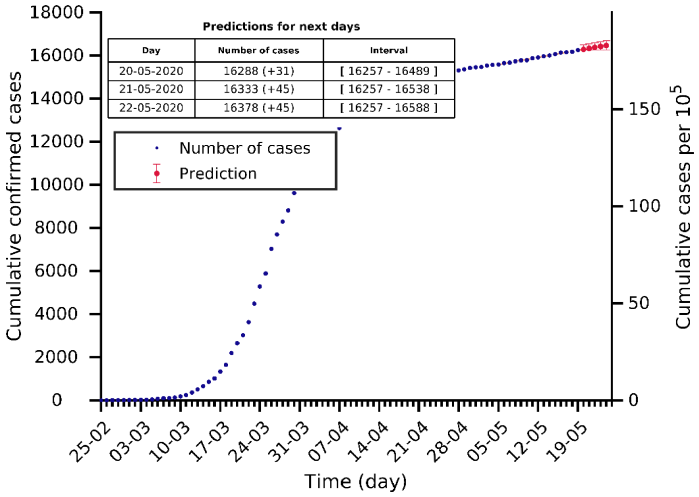


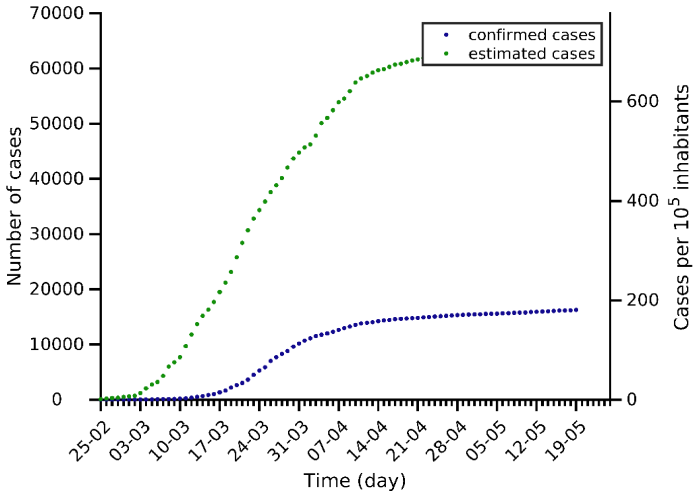


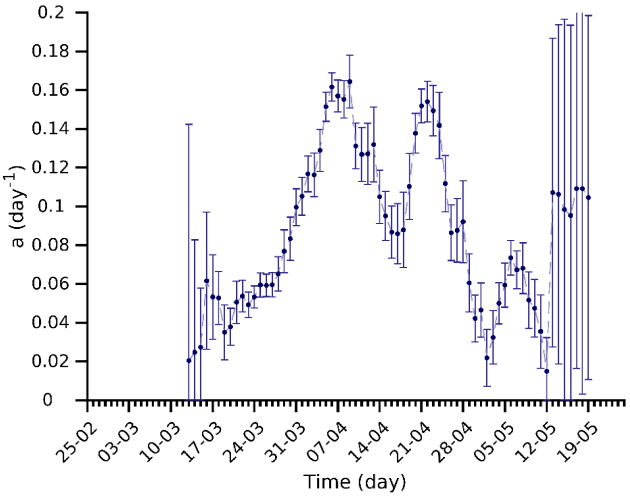


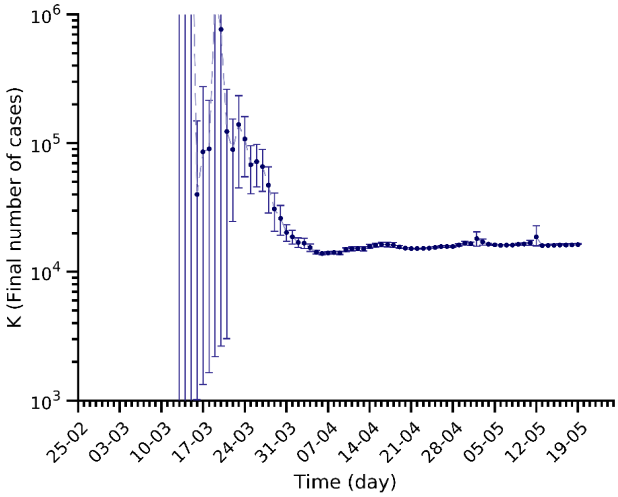


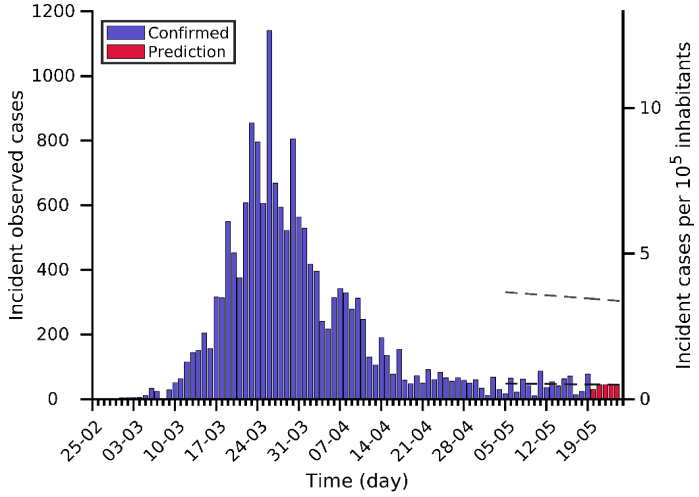
Austria 19-05-2020. Population: 9.0M. Current cumulated incidence: 181/10^5

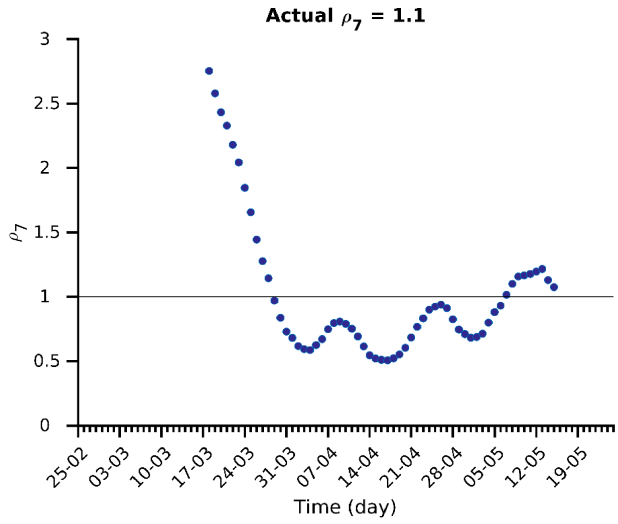


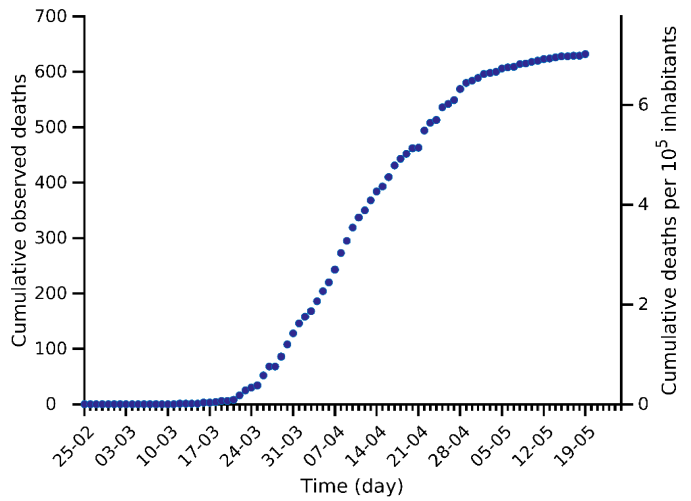


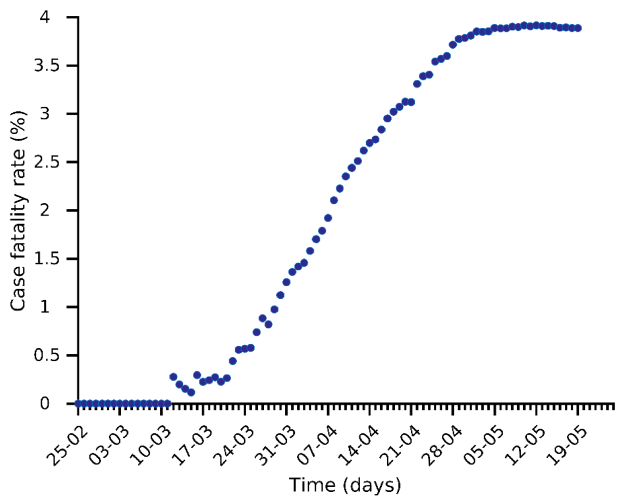




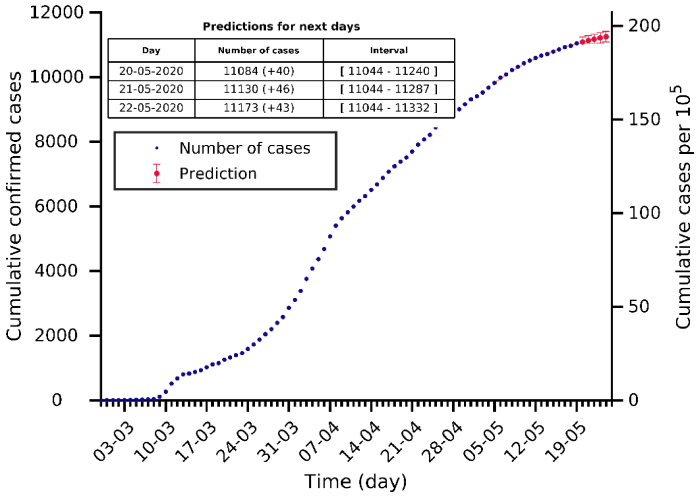


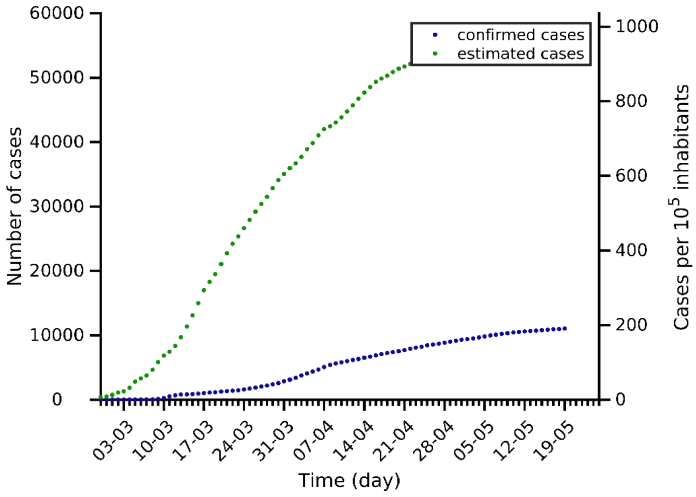


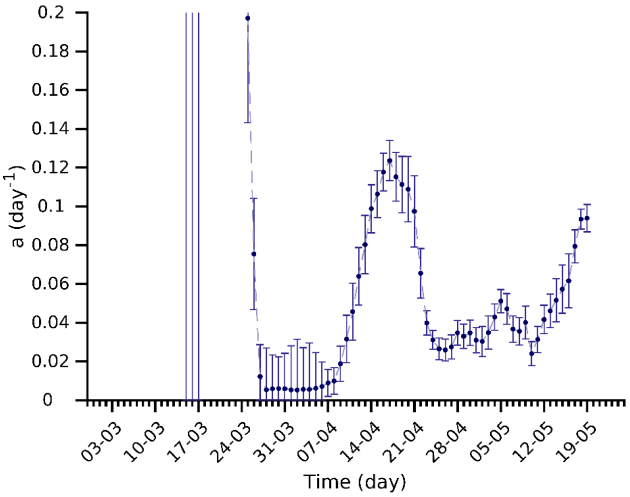


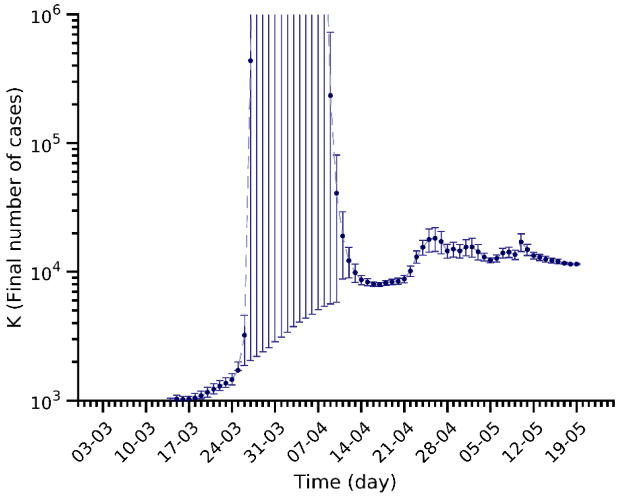


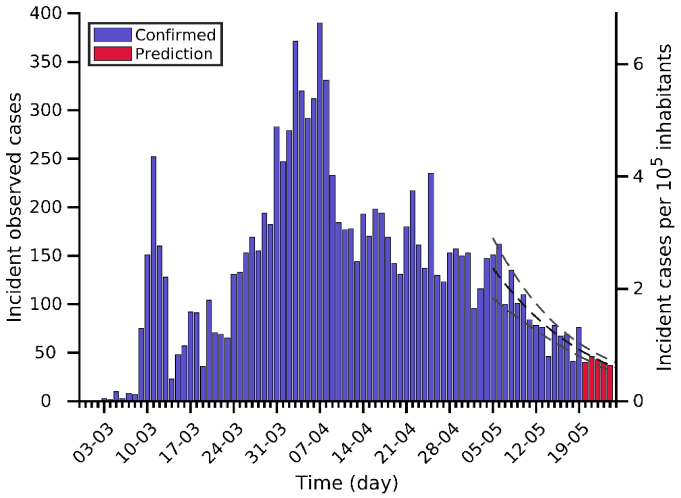
Denmark 19-05-2020. Population: 5.8M. Current cumulated incidence: 191/10^5

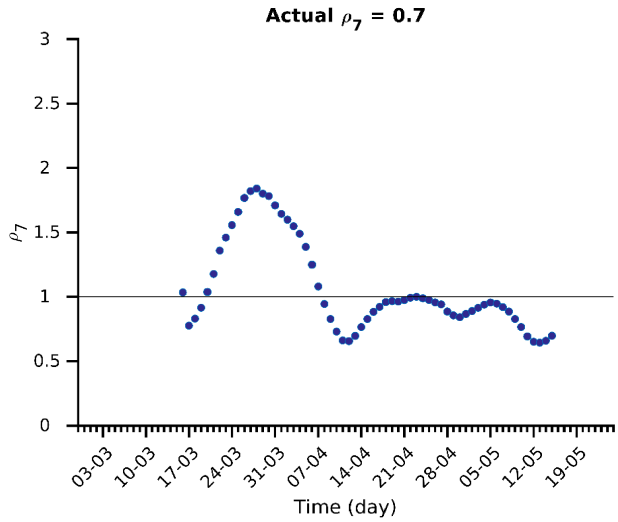


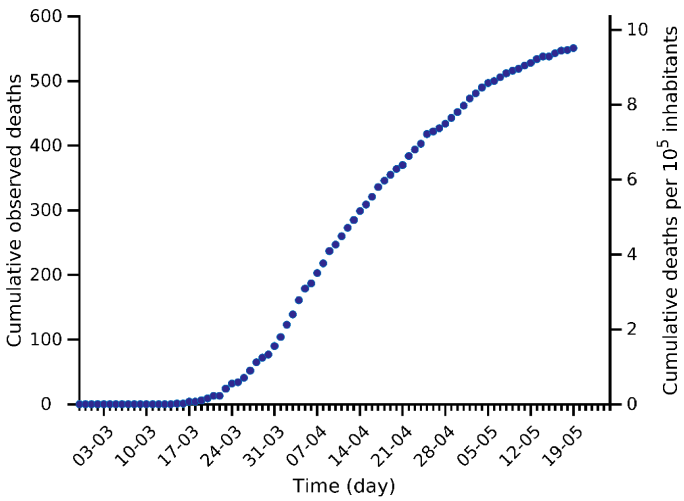


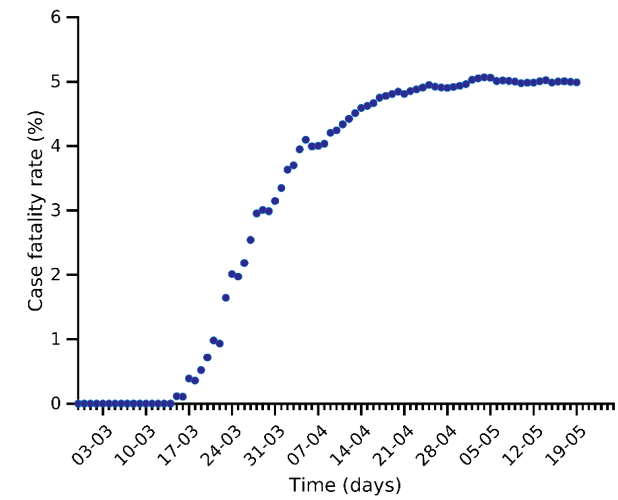




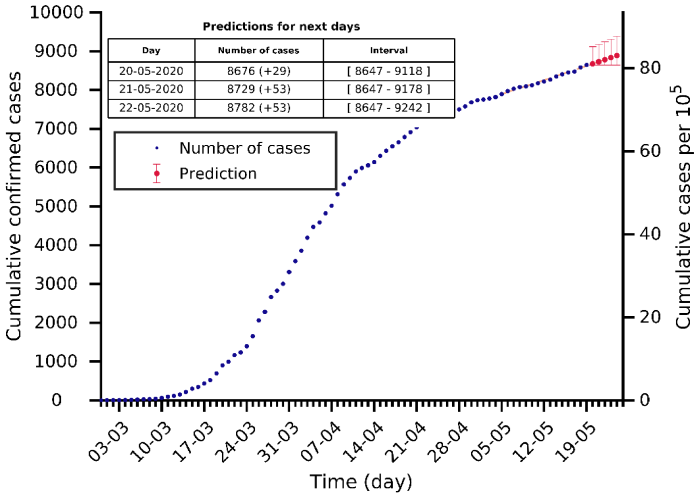


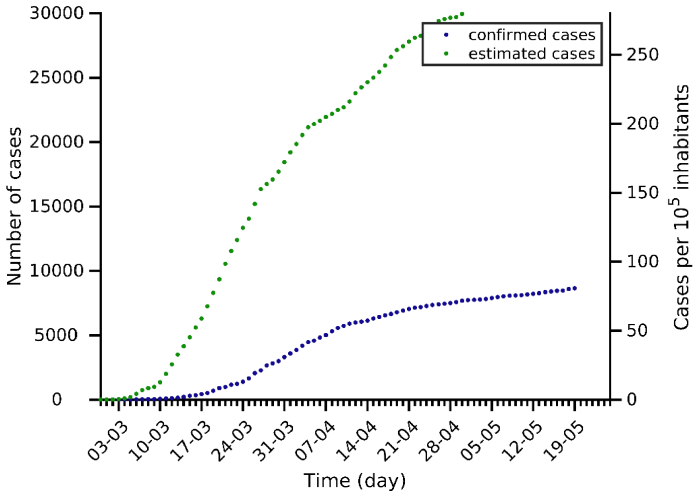


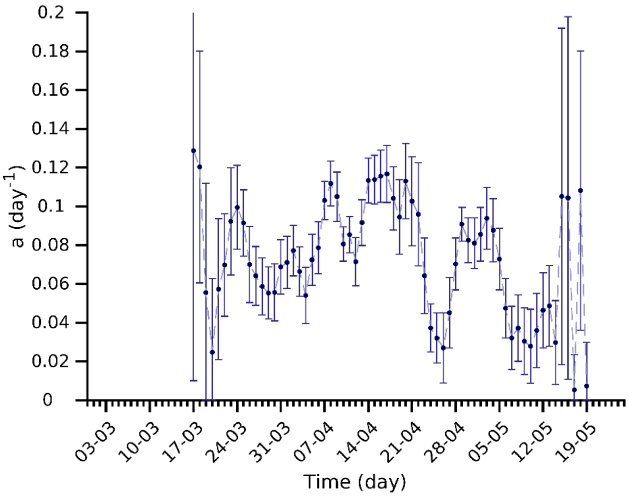


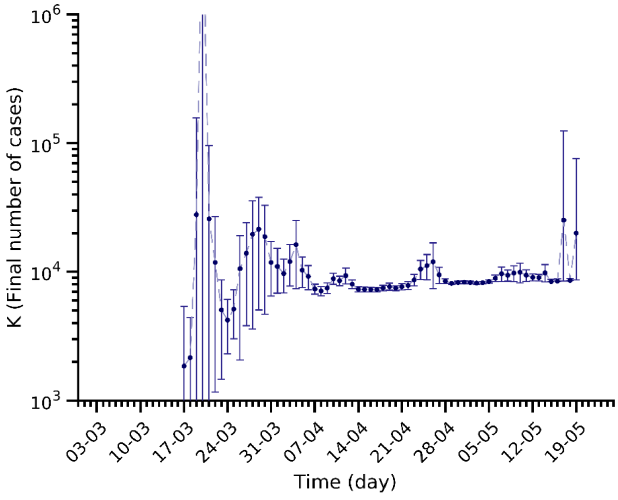


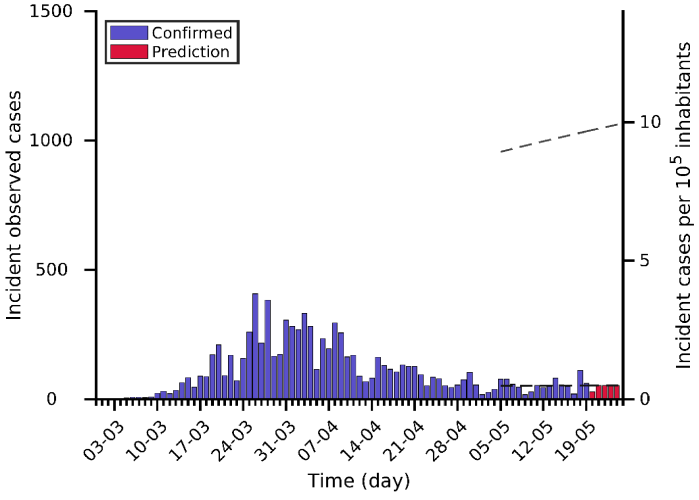
Czech Rep 19-05-2020. Population: 10.7M. Current cumulated incidence: 81/10^5

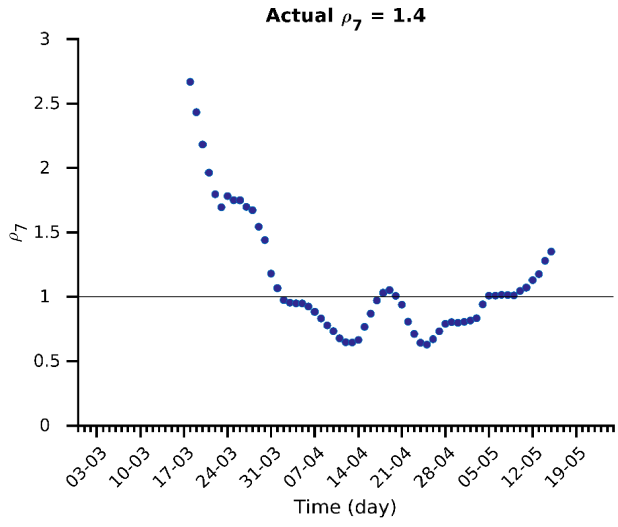


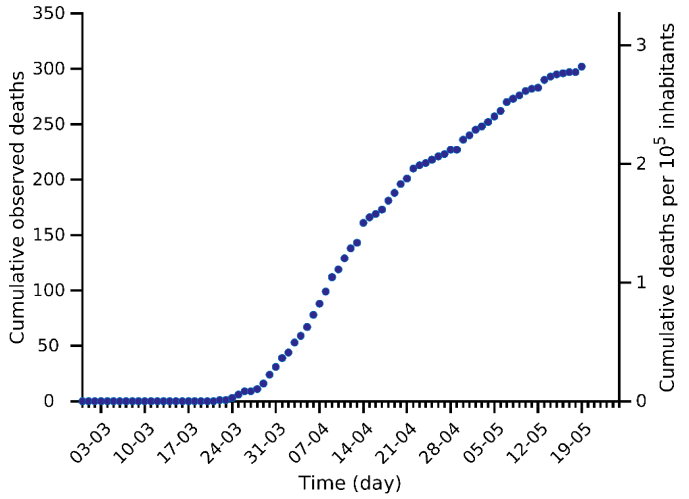


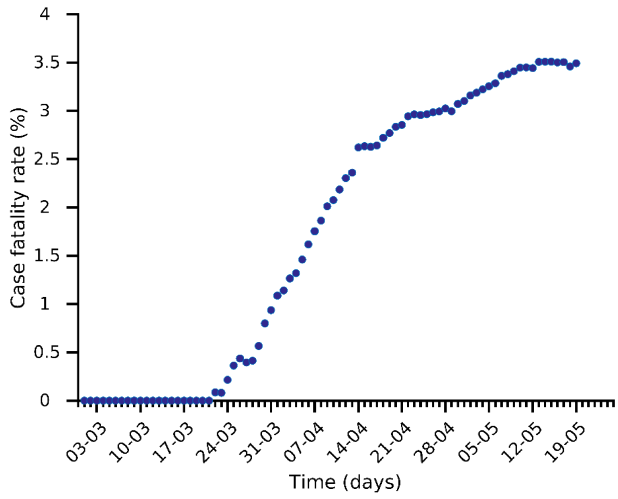




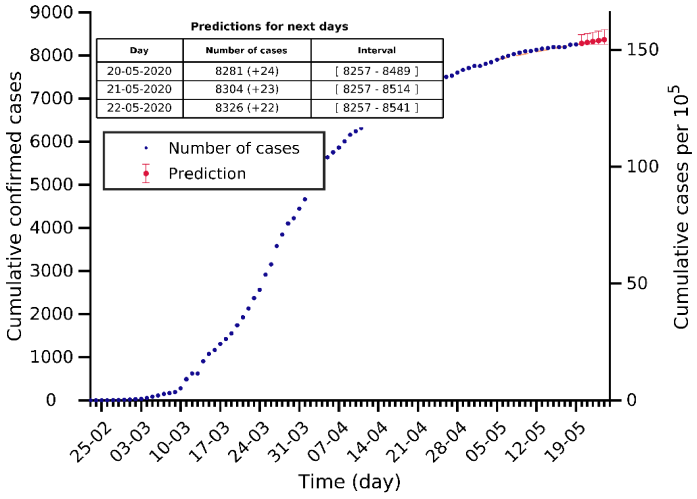


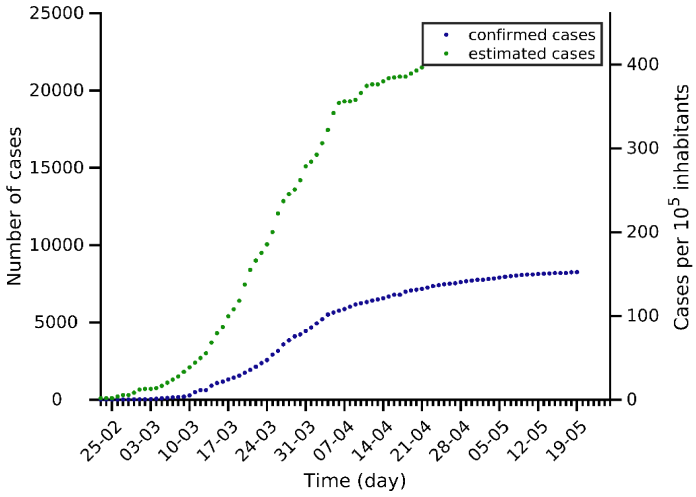


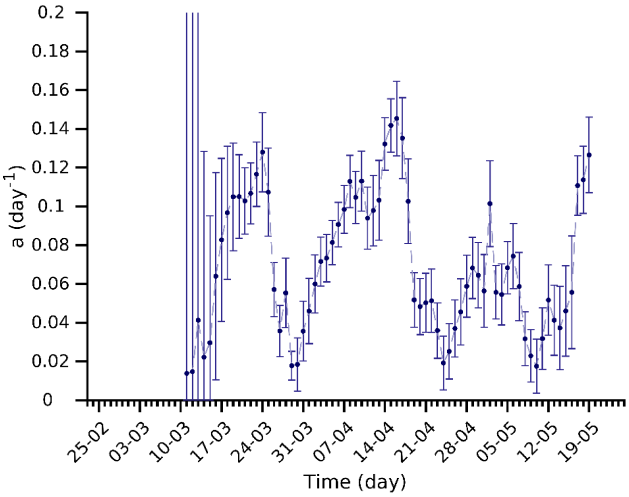


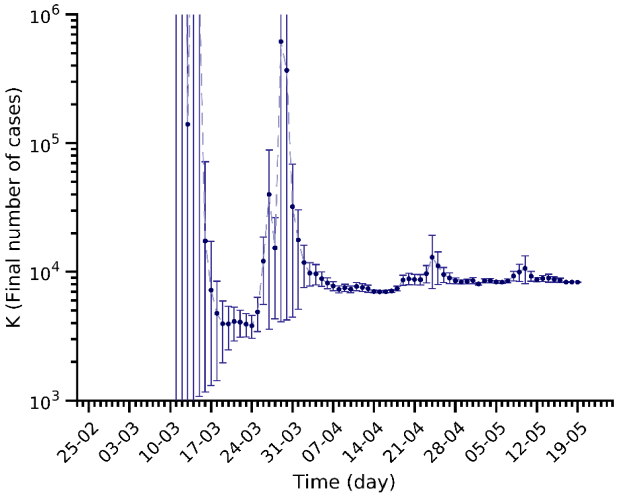


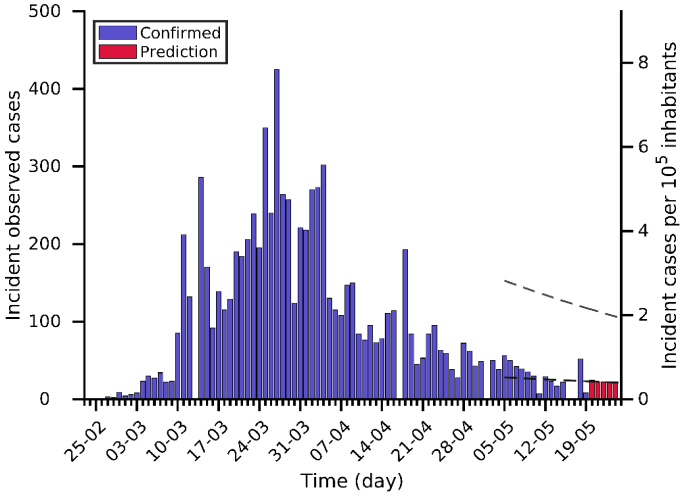
Norway 19-05-2020. Population: 5.4M. Current cumulated incidence: 152/10^5

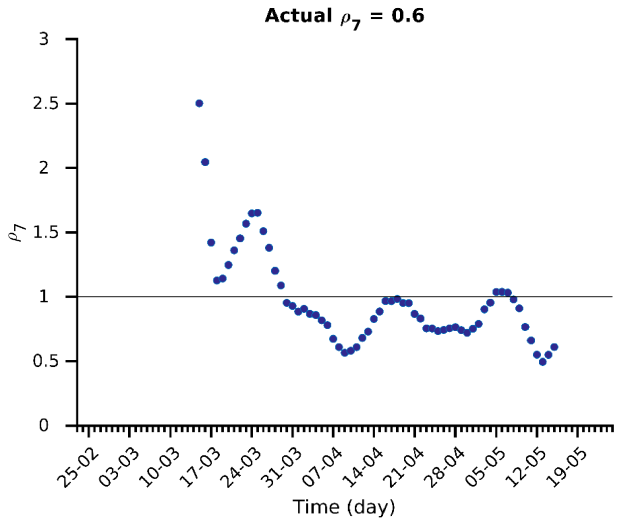


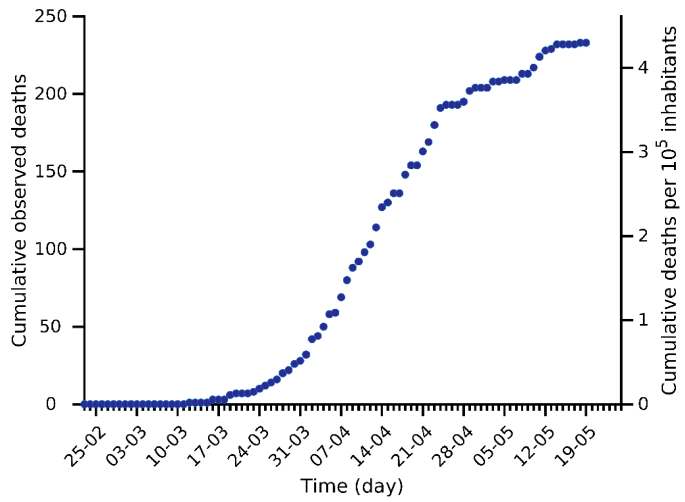


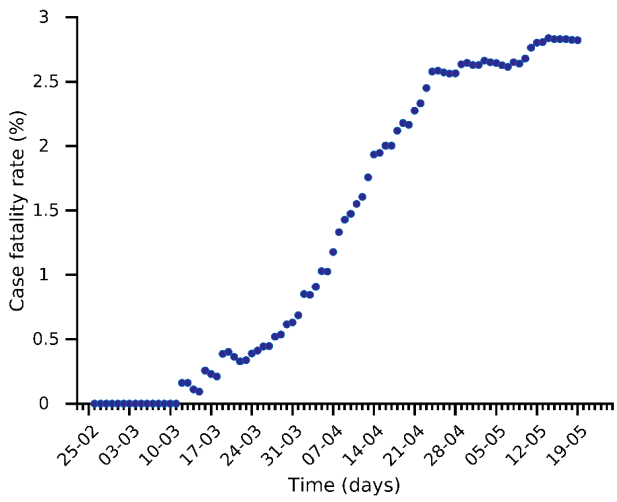




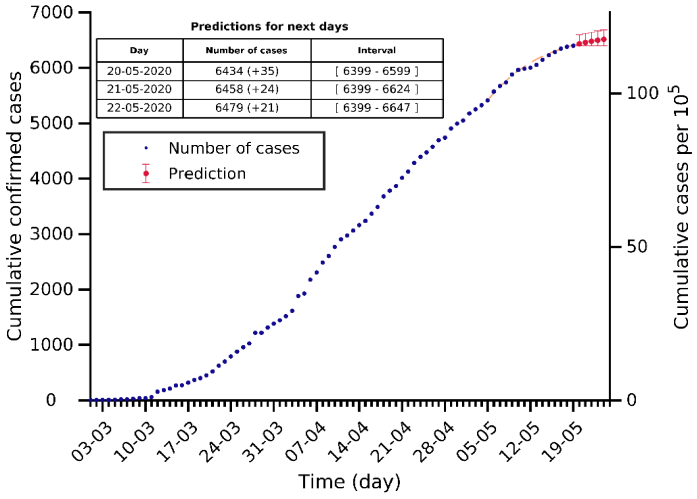


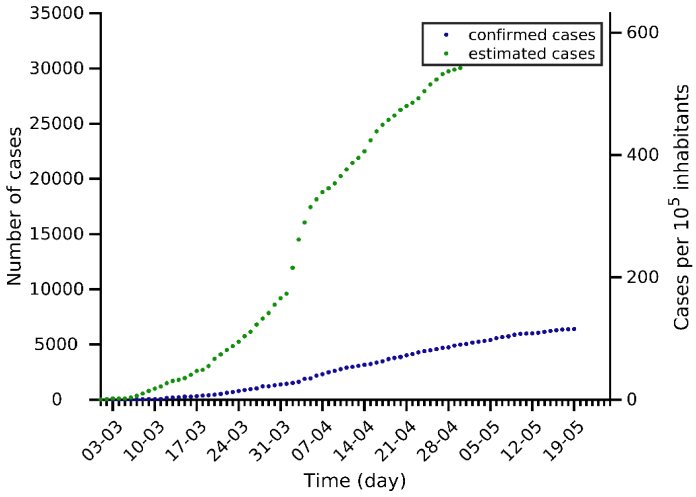


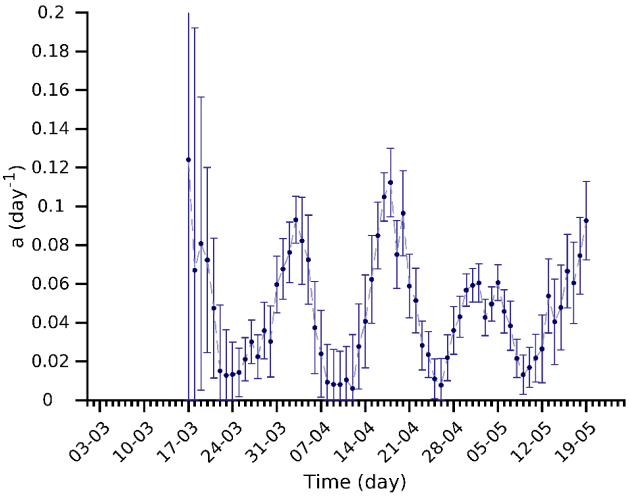


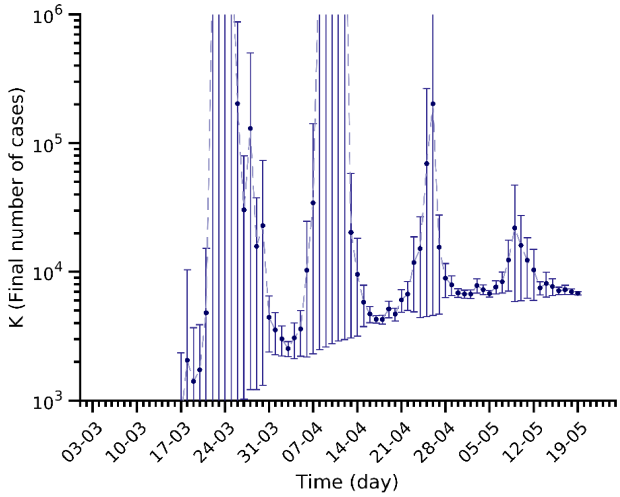


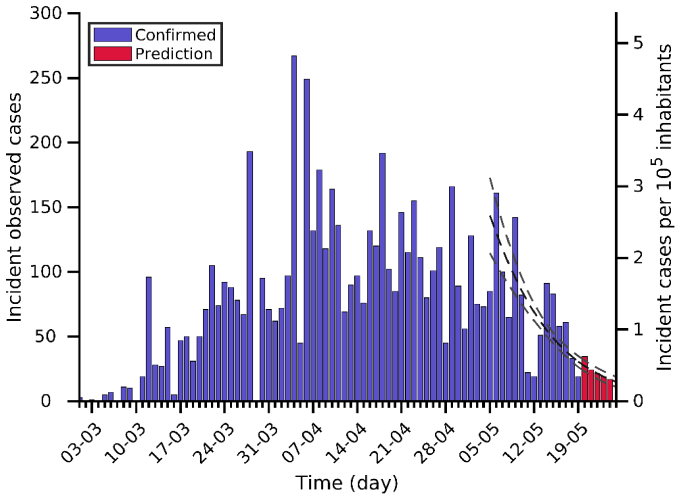
Finland 19-05-2020. Population: 5.5M. Current cumulated incidence: 115/10⁵

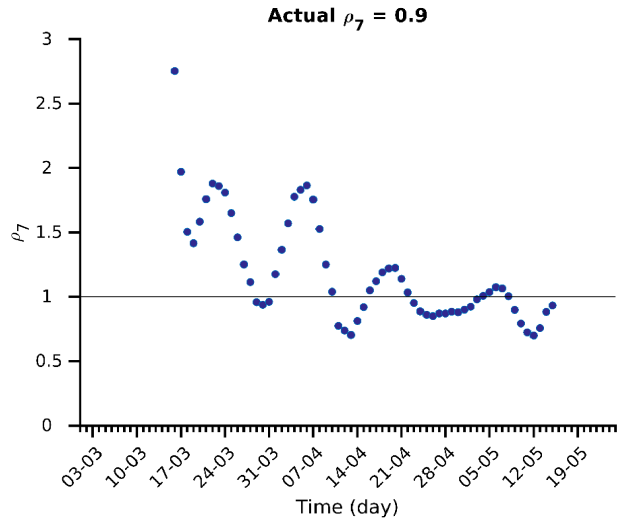


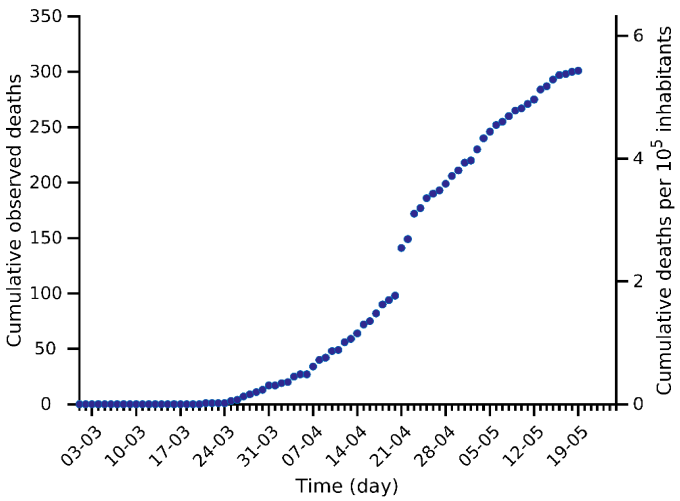


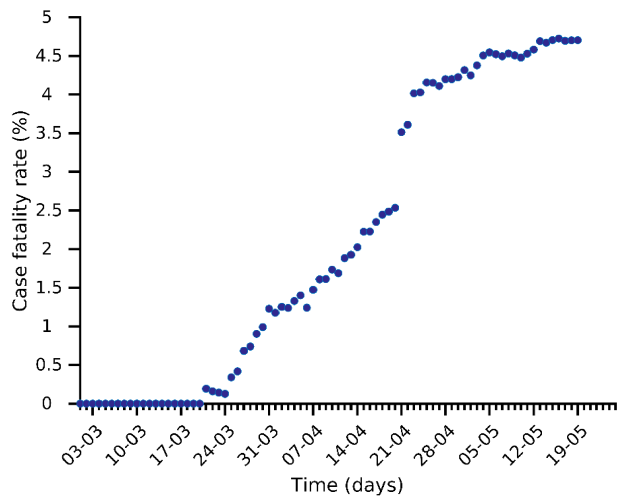




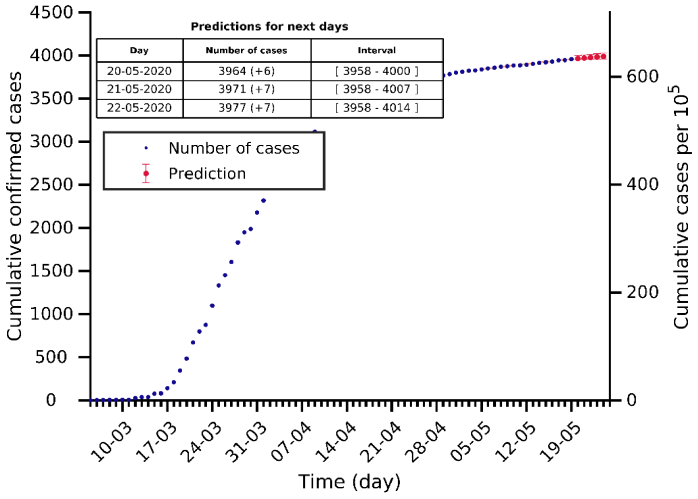


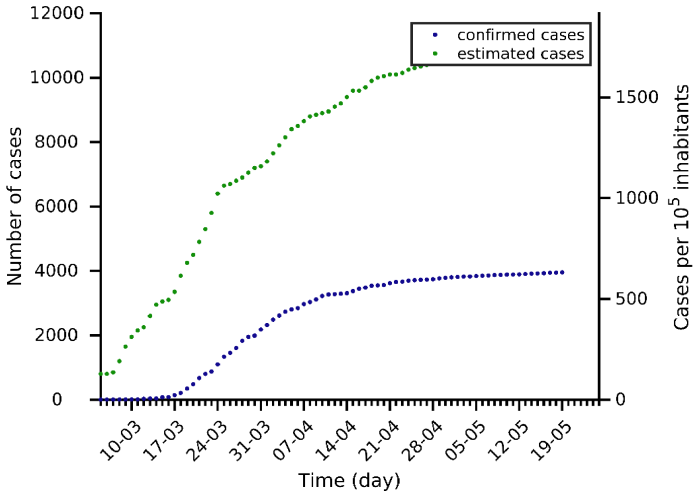


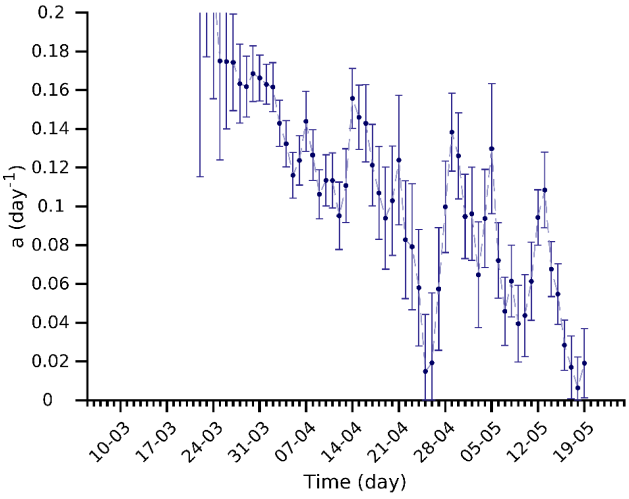


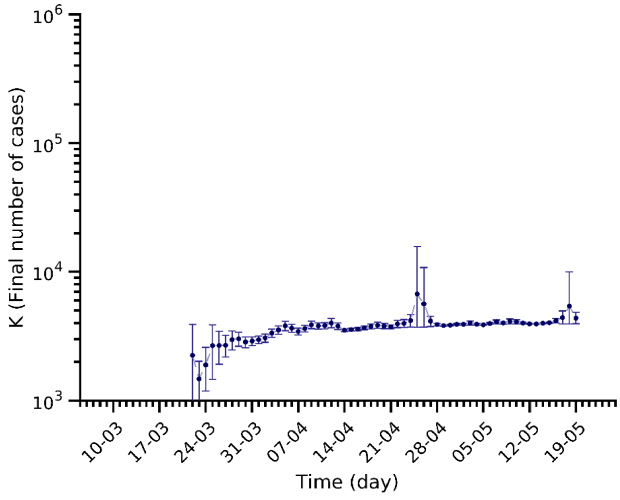


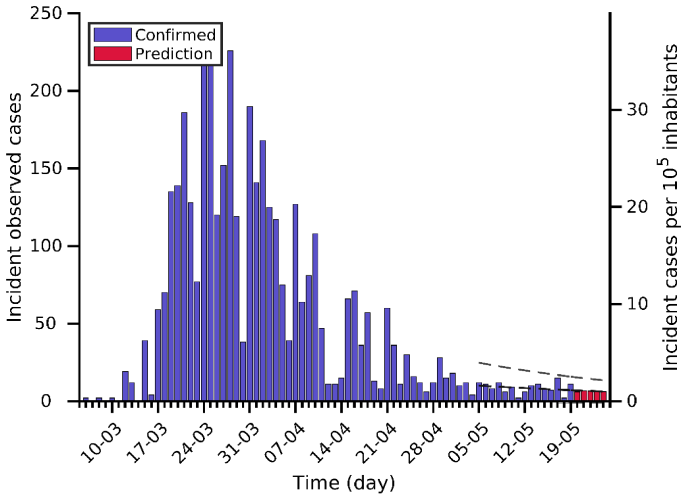
Luxembourg 19-05-2020. Population: 0.6M. Current cumulated incidence: 632/10^5

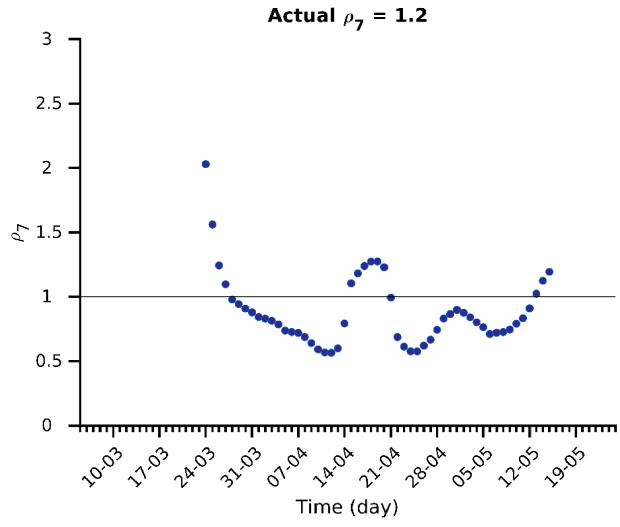


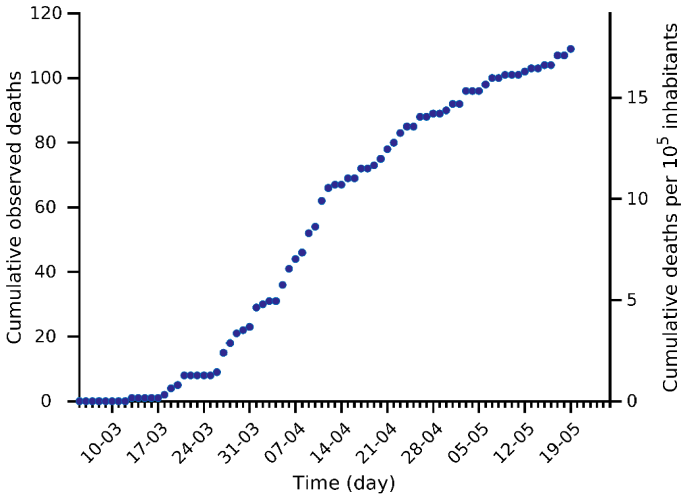


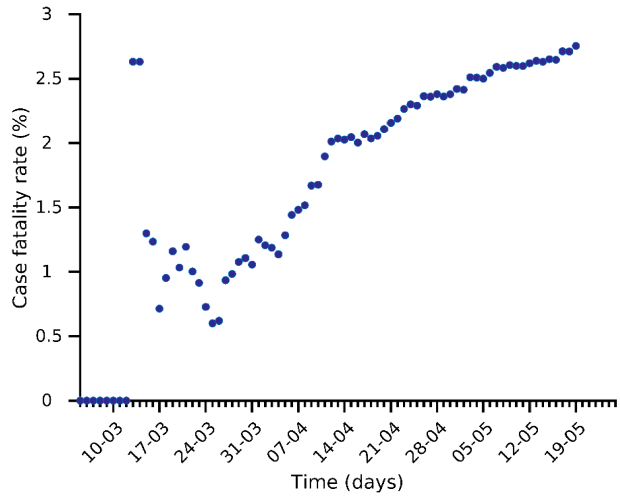




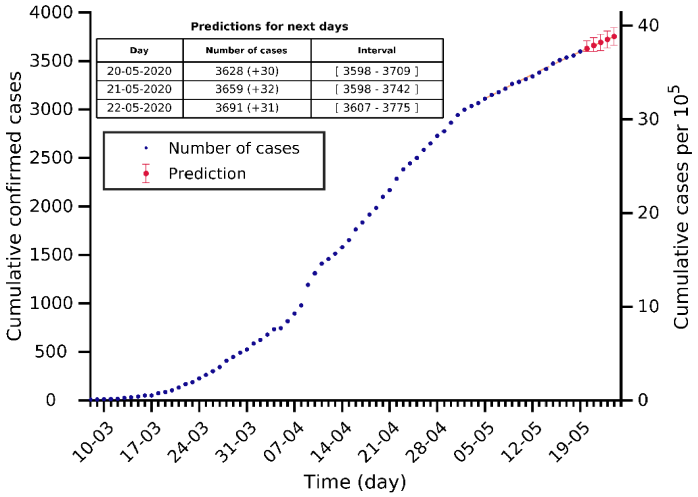


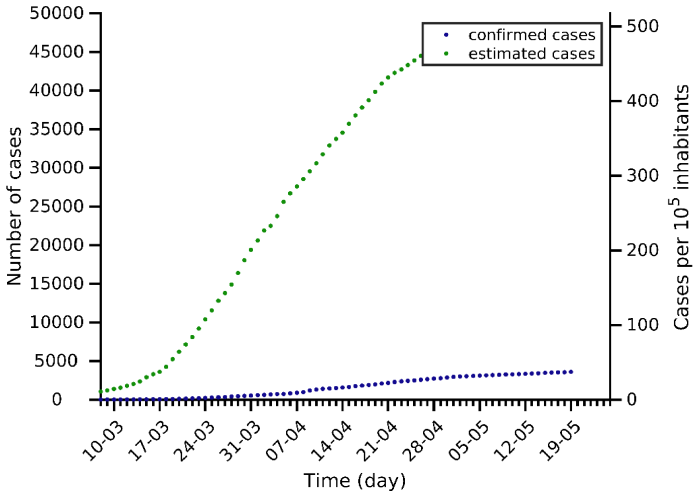


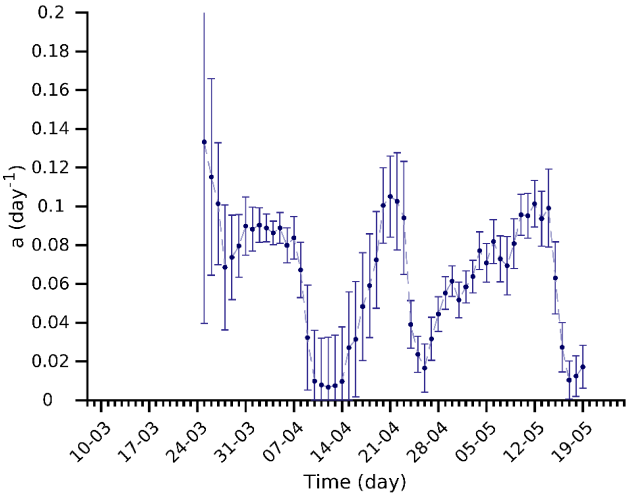


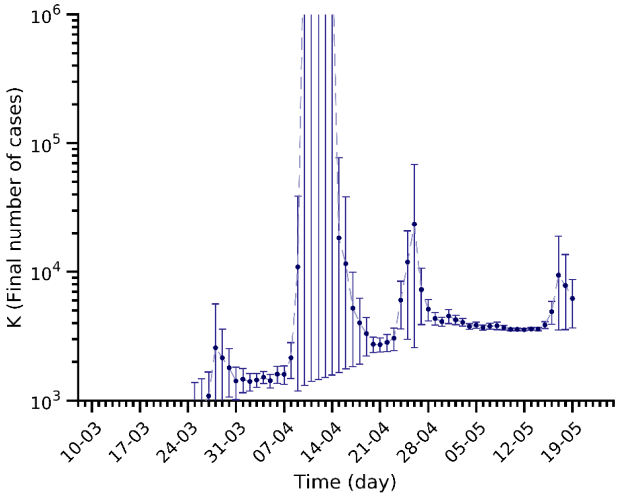


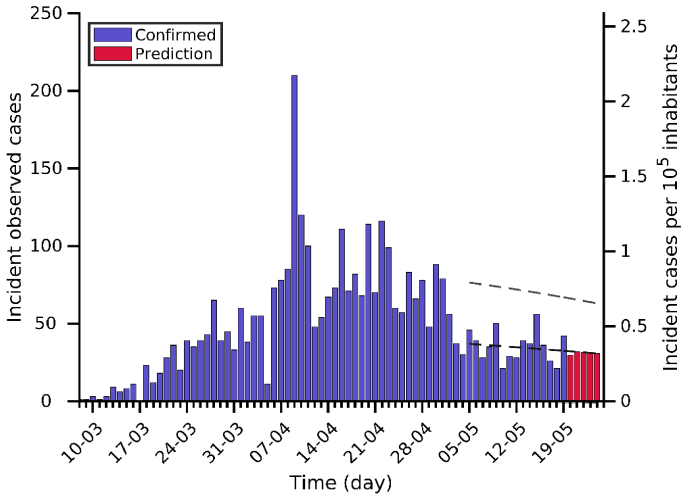
Hungary 19-05-2020. Population: 9.7M. Current cumulated incidence: $37/10^5$

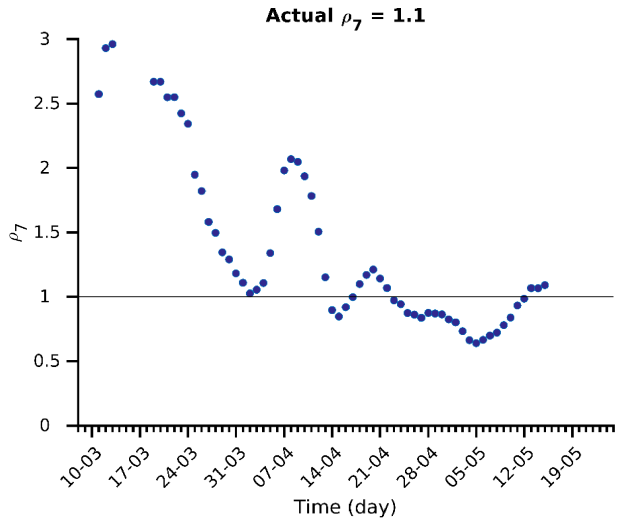


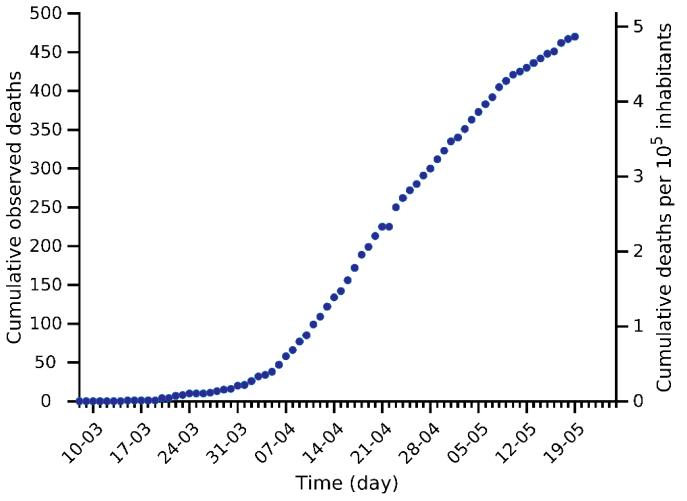


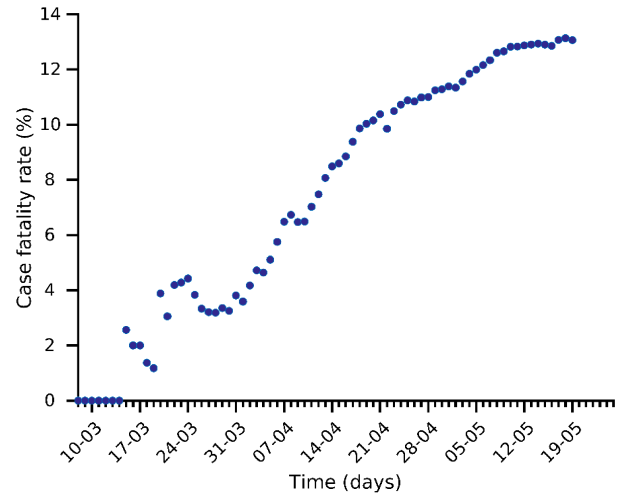




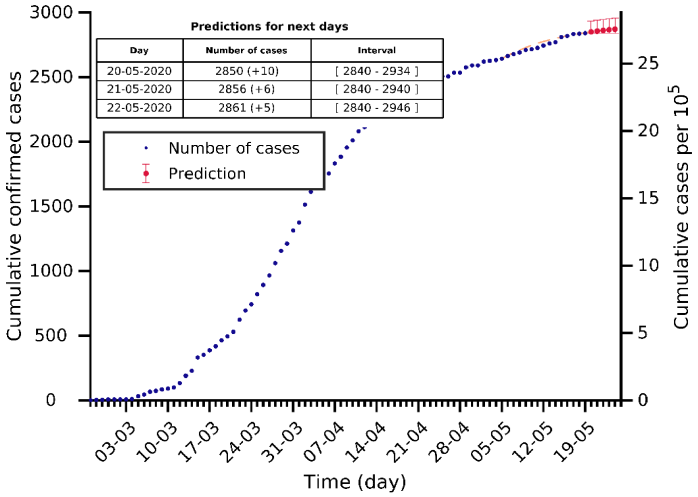


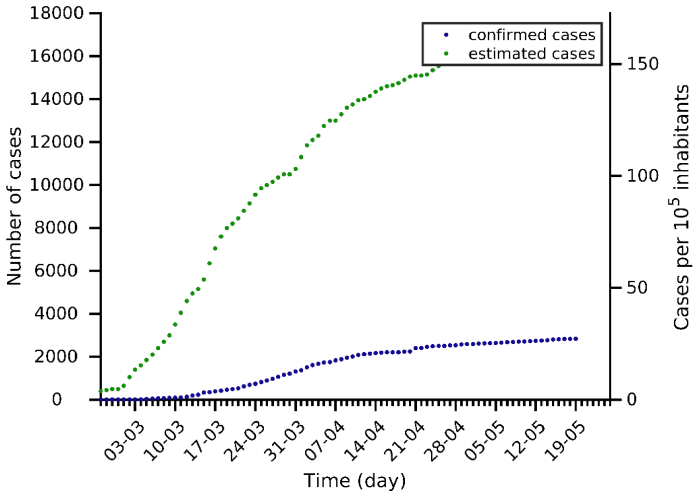


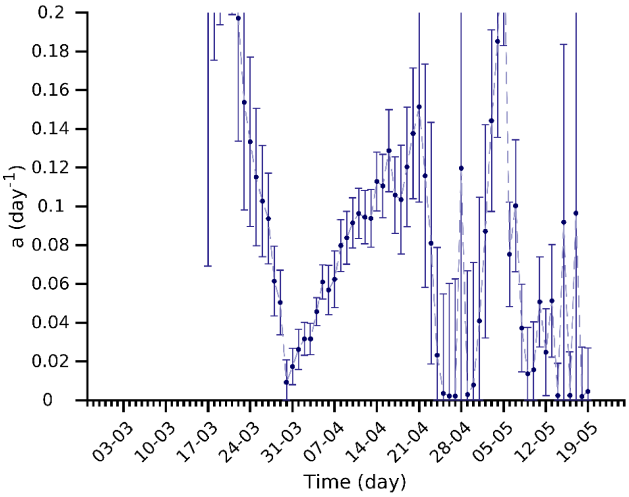


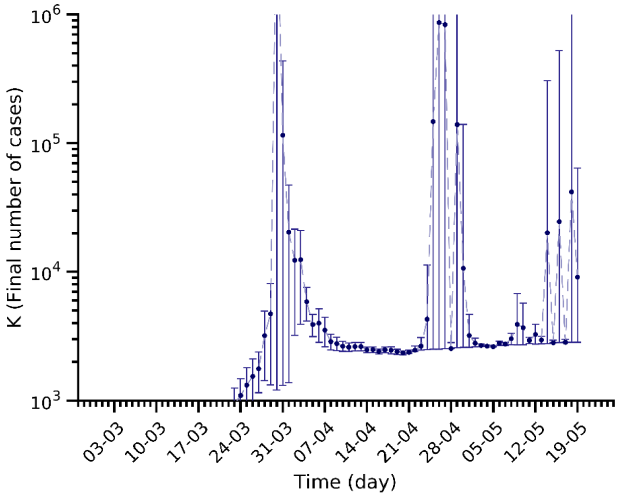


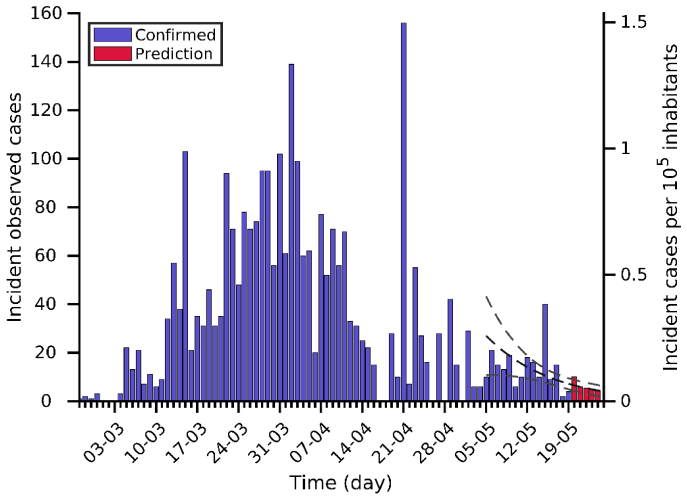
Greece 19-05-2020. Population: 10.4M. Current cumulated incidence: 27/10^5

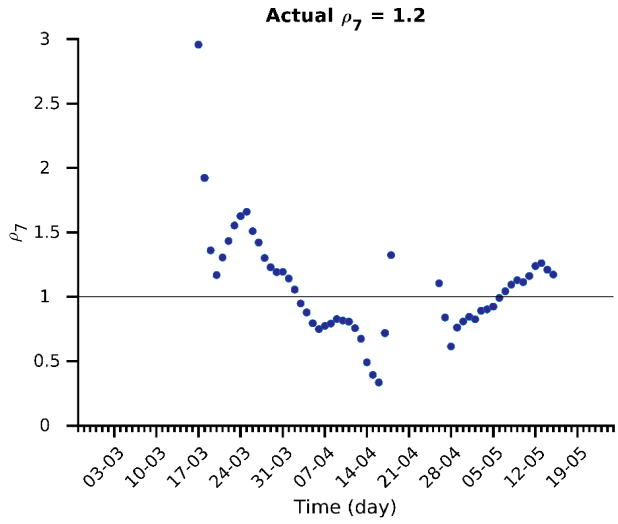


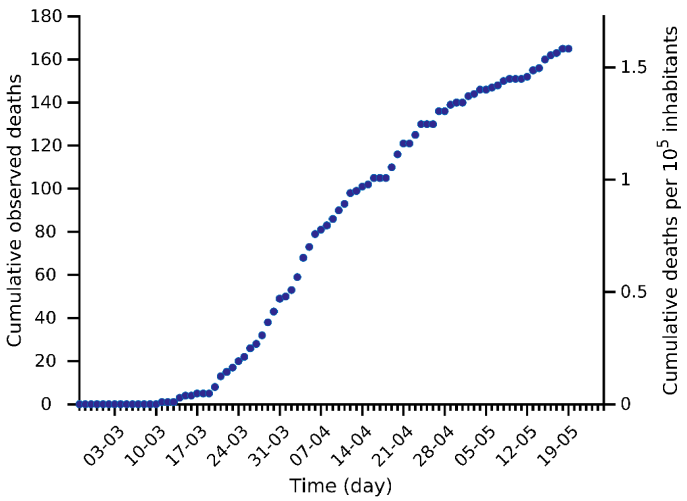


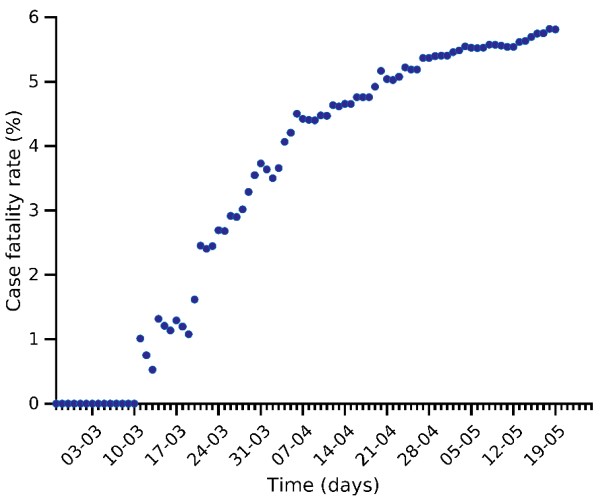




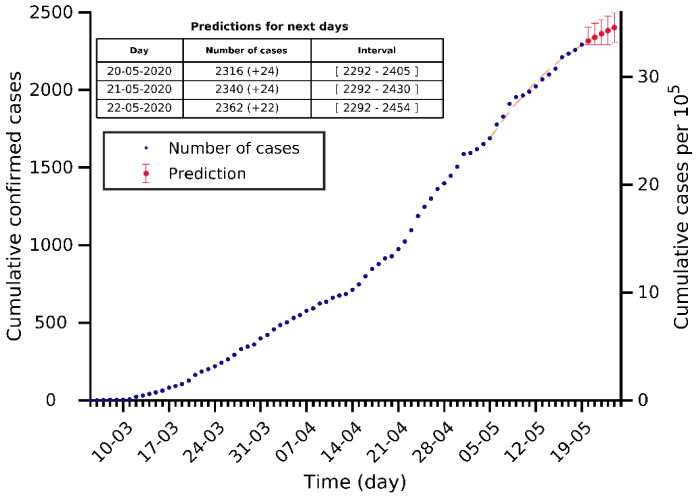


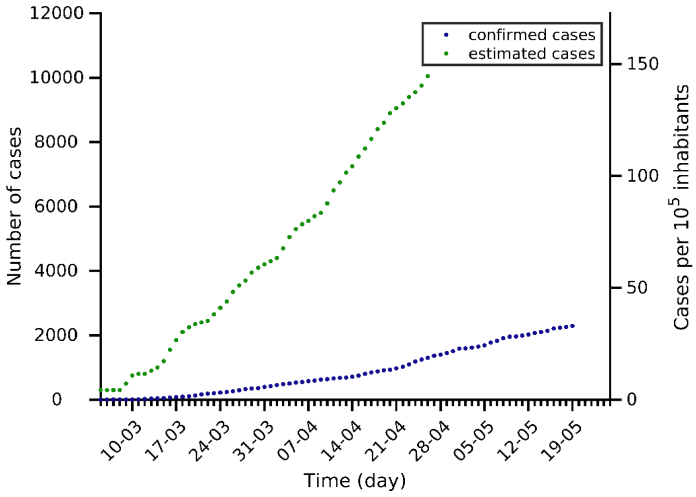


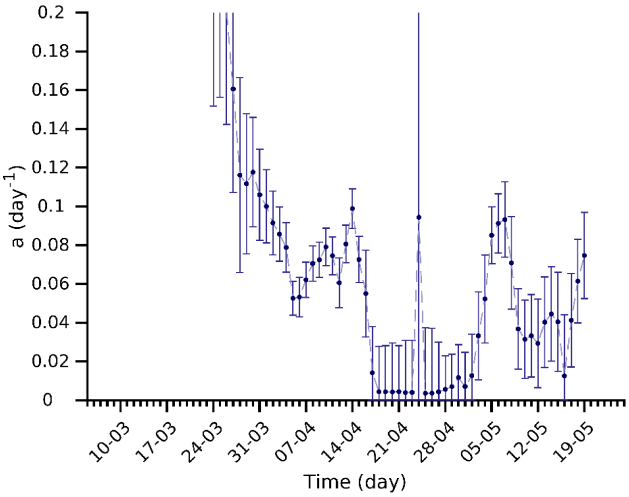


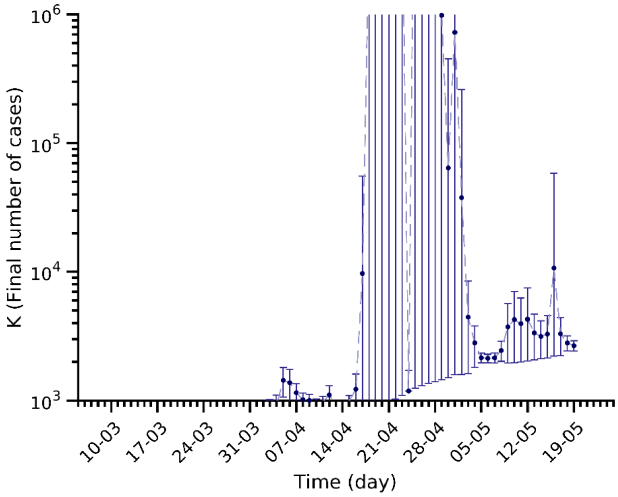


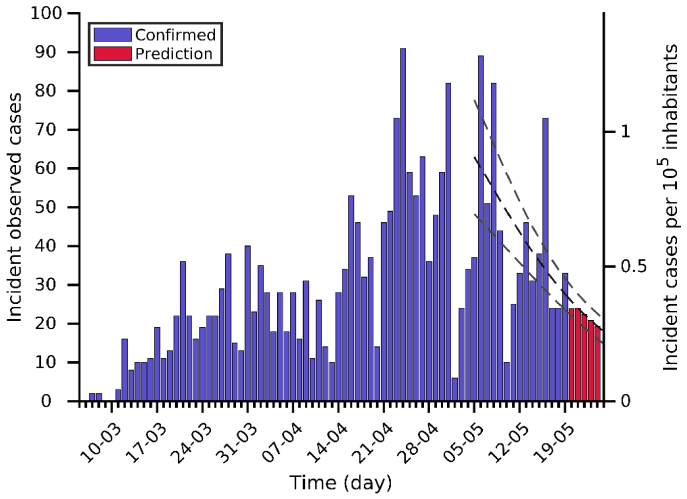
Bulgaria 19-05-2020. Population: 6.9M. Current cumulated incidence: $33/10^5$

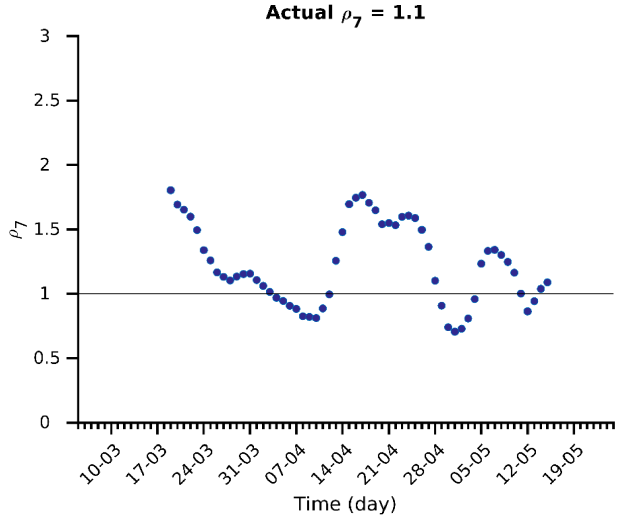


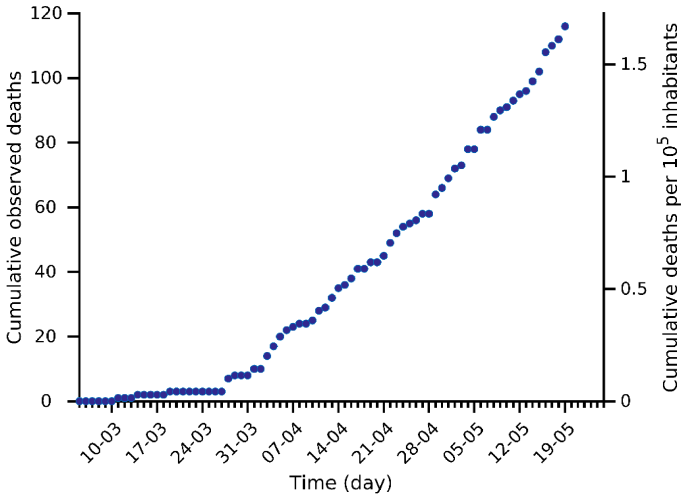


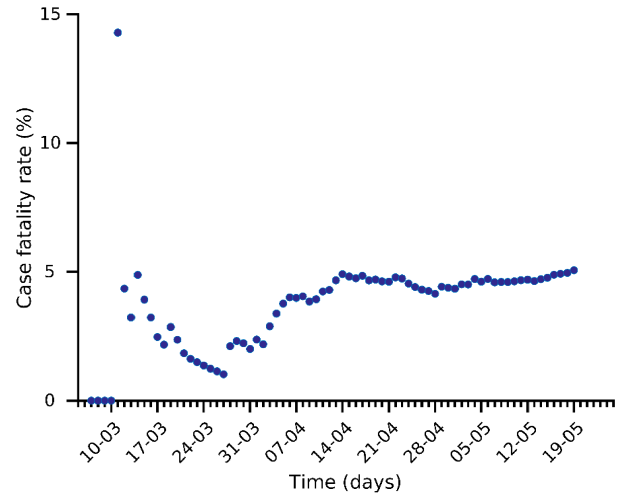




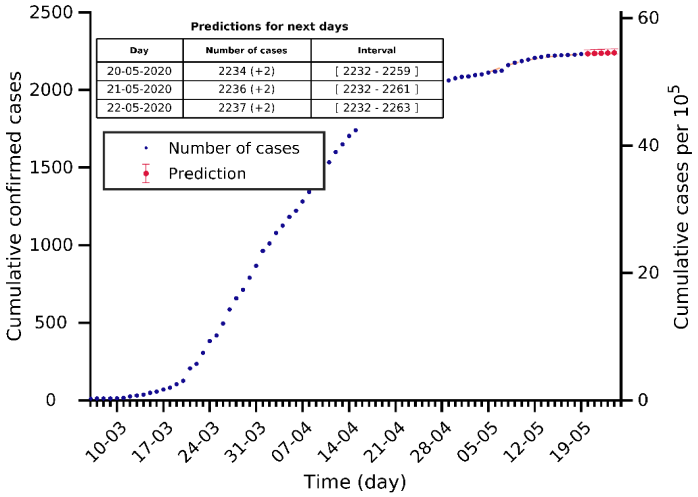


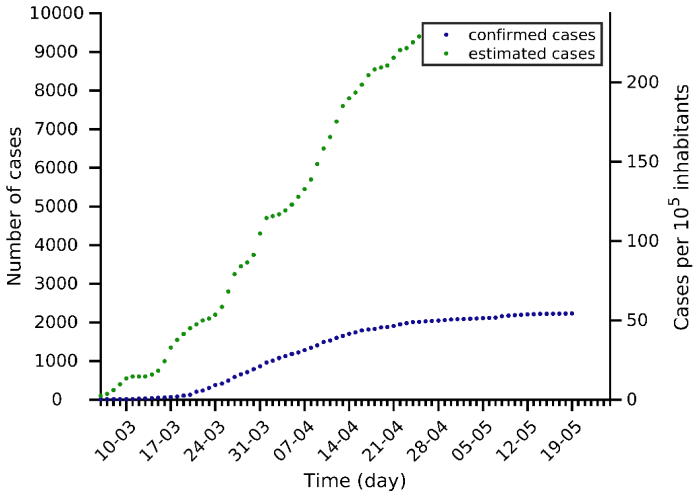


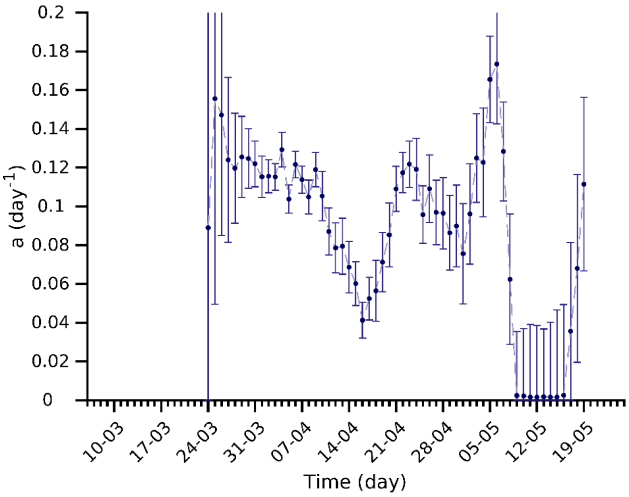


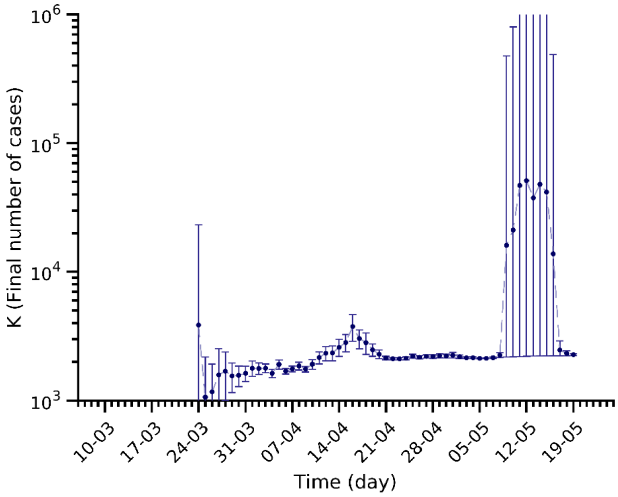


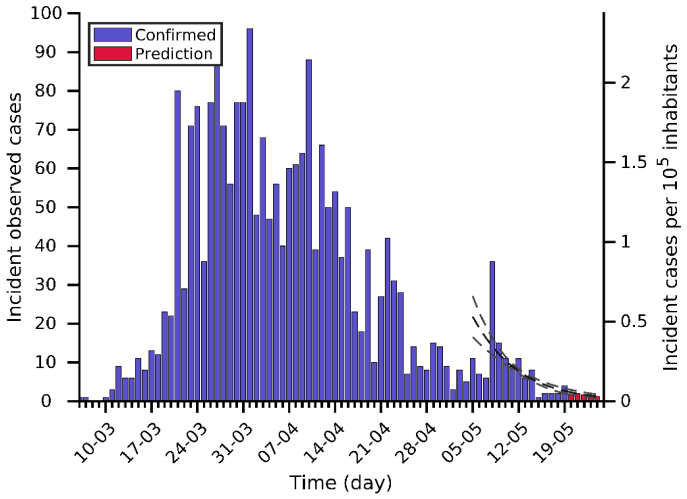
Croatia 19-05-2020. Population: 4.1M. Current cumulated incidence: $54/10^5$

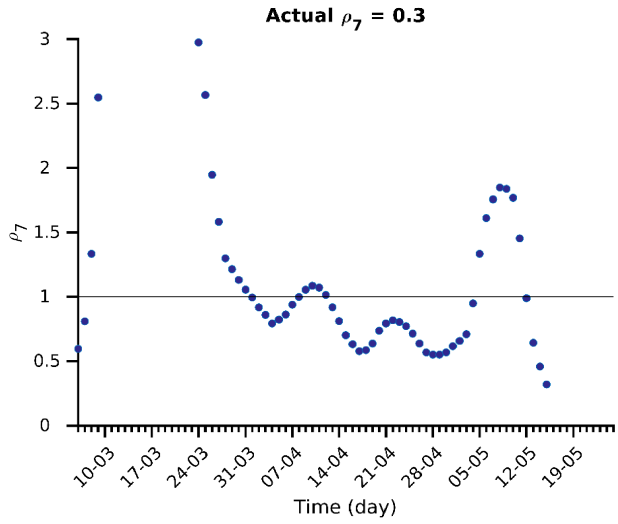


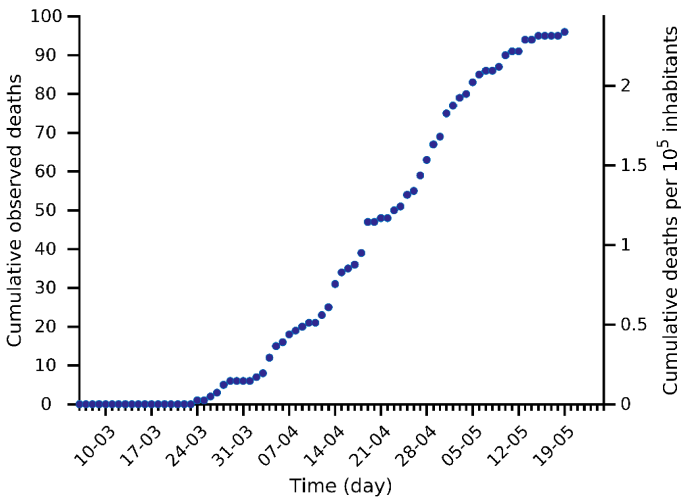


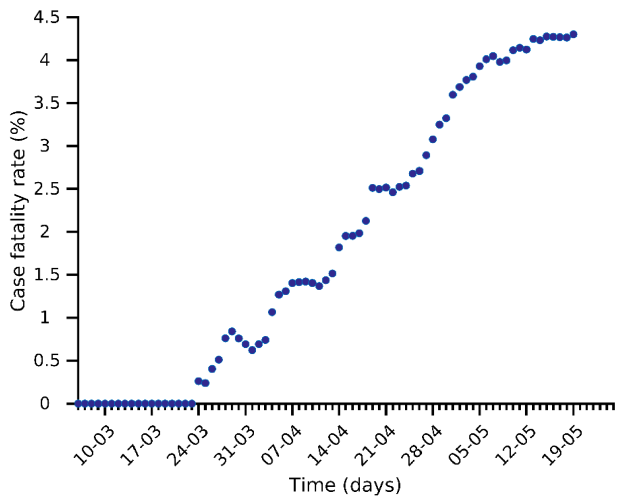




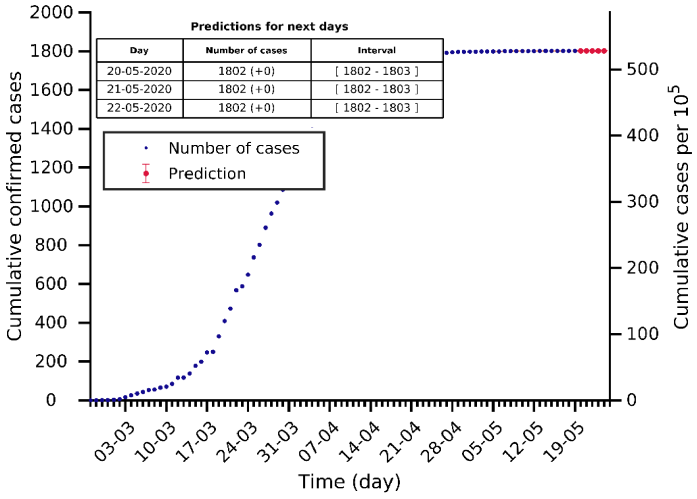


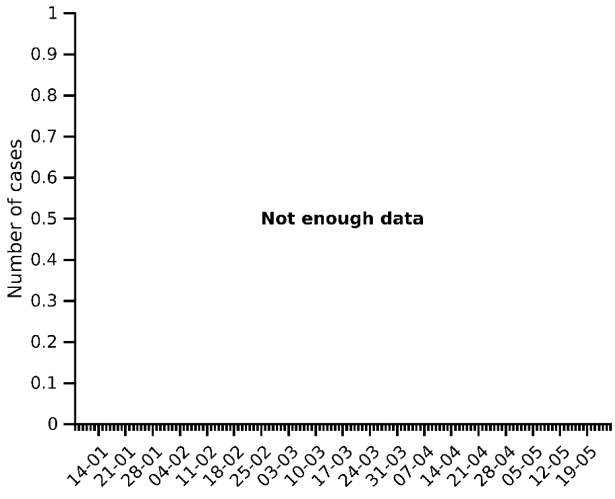


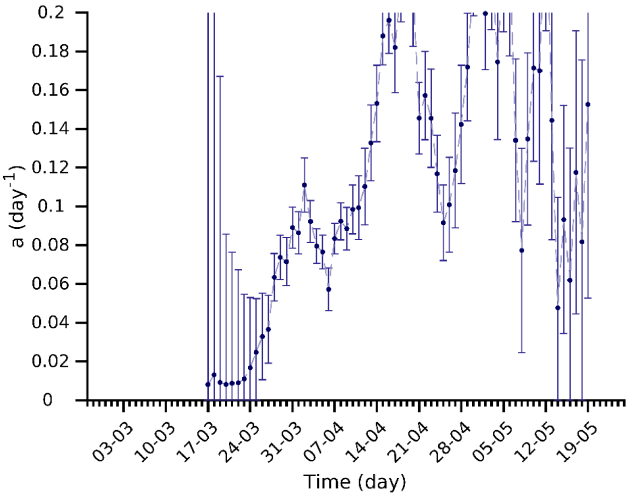


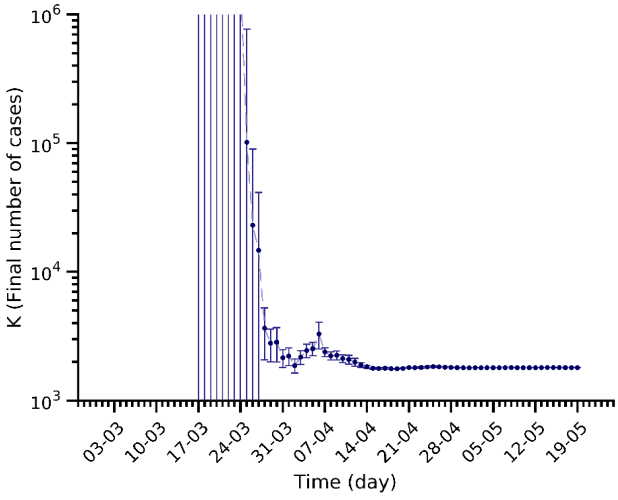


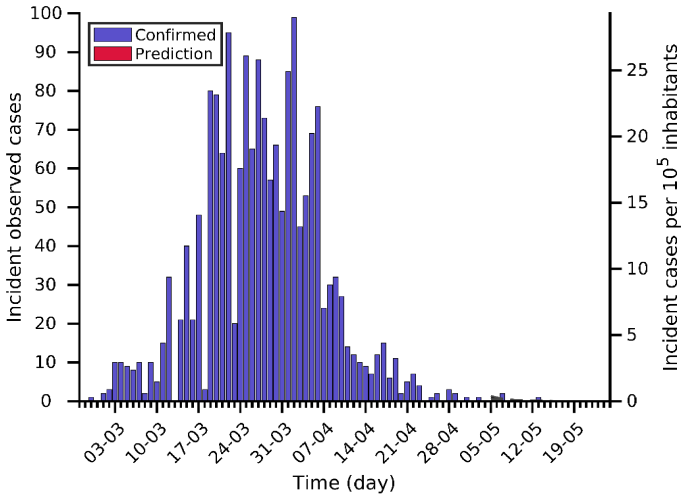
Iceland 19-05-2020. Population: 0.3M. Current cumulated incidence: $528/10^5$

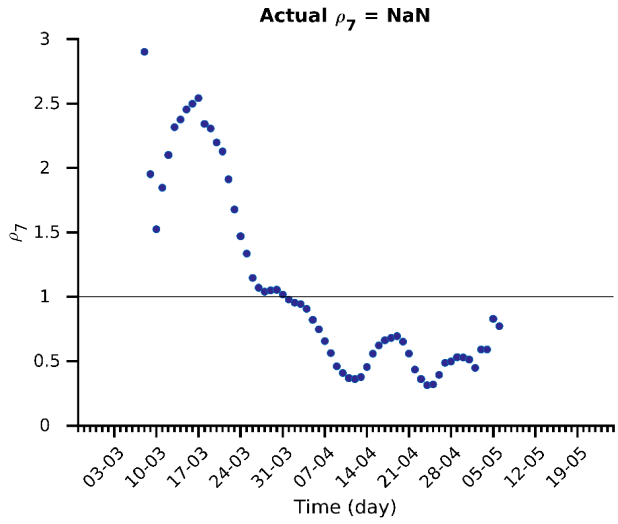


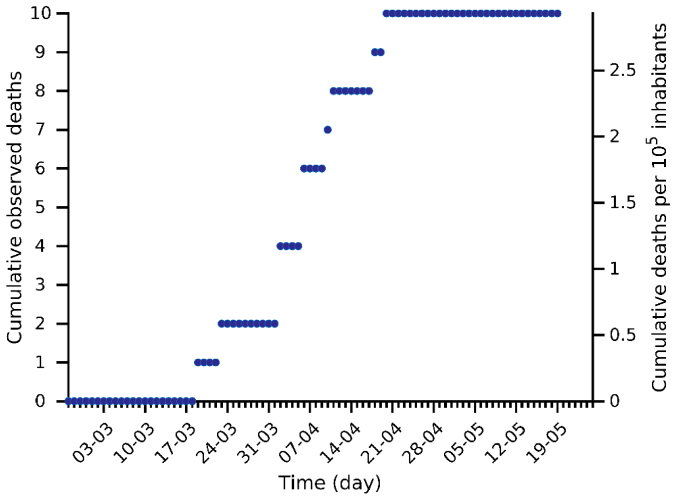


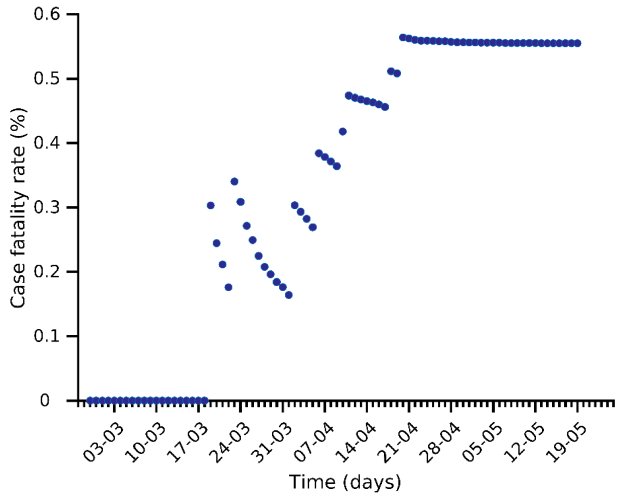




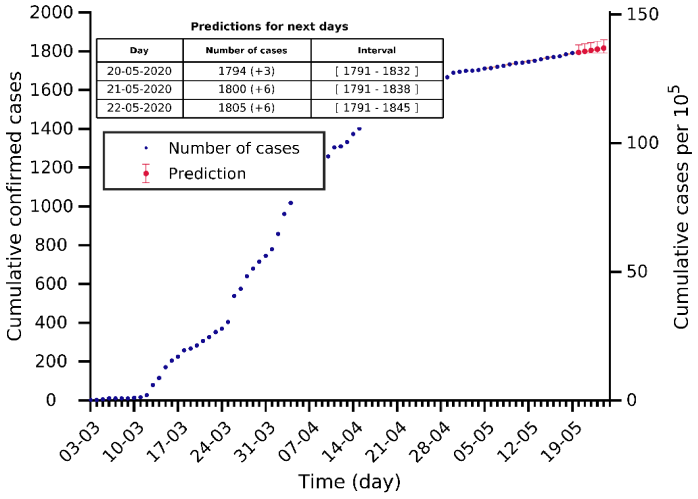


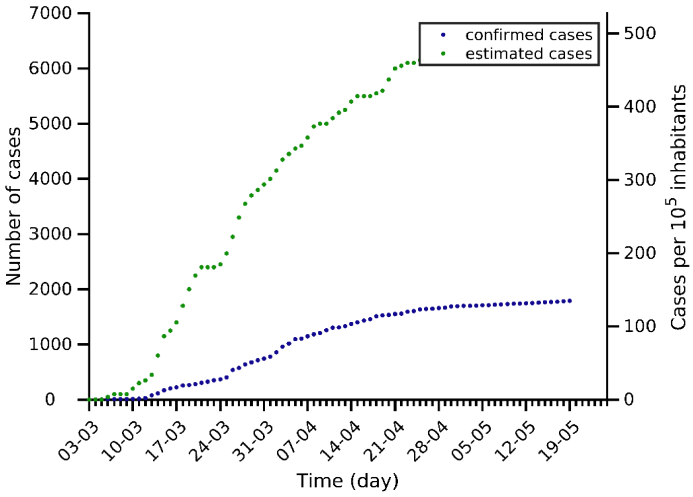


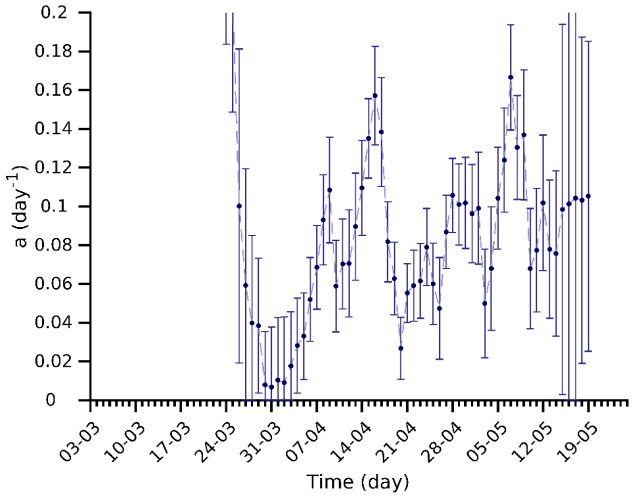


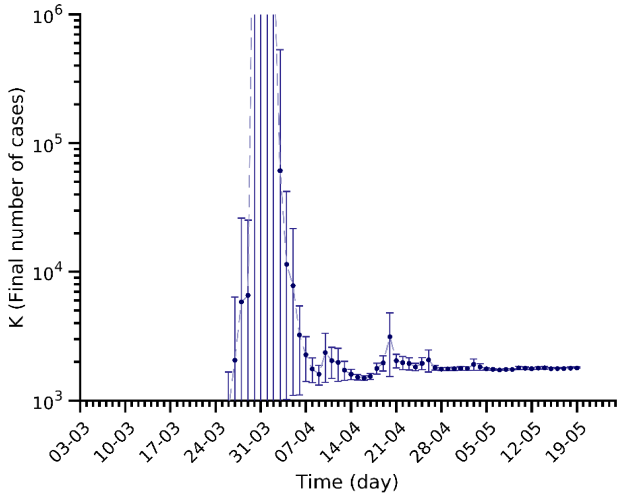


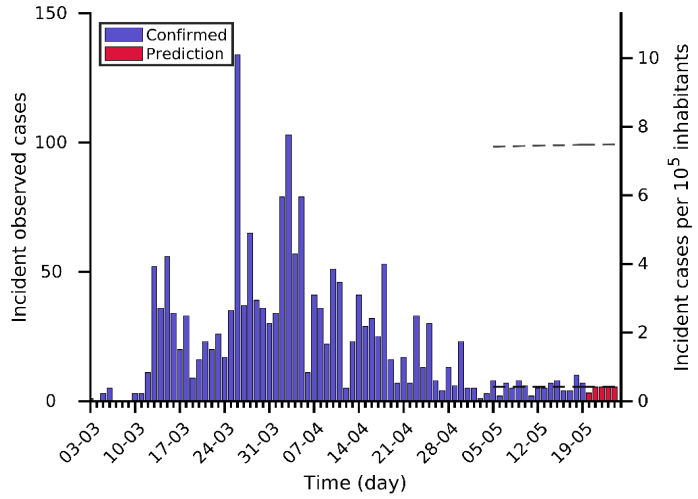
Estonia 19-05-2020. Population: 1.3M. Current cumulated incidence: $135/10^5$

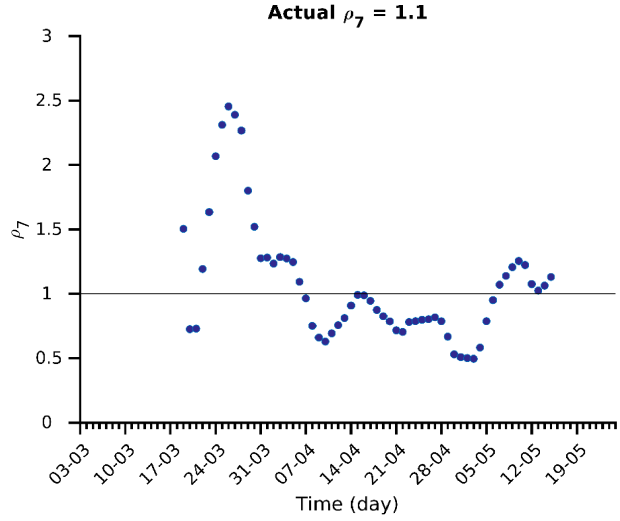


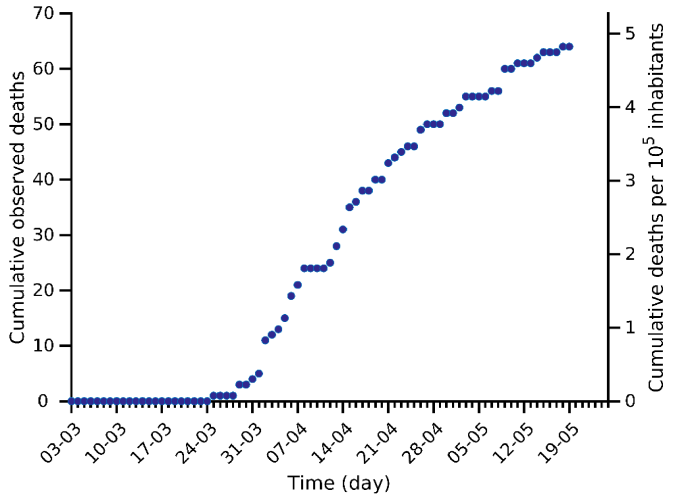


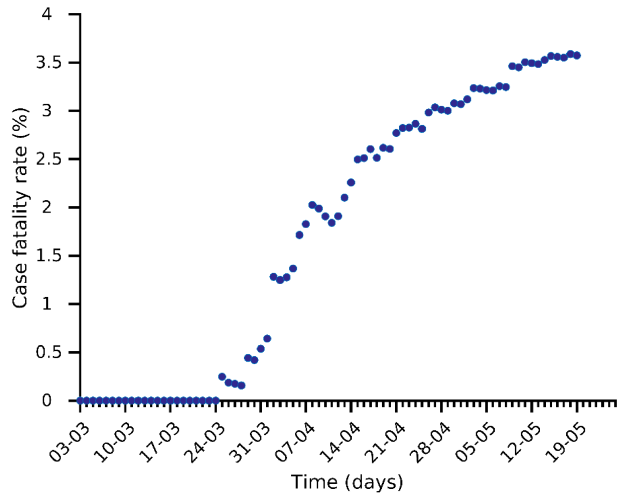




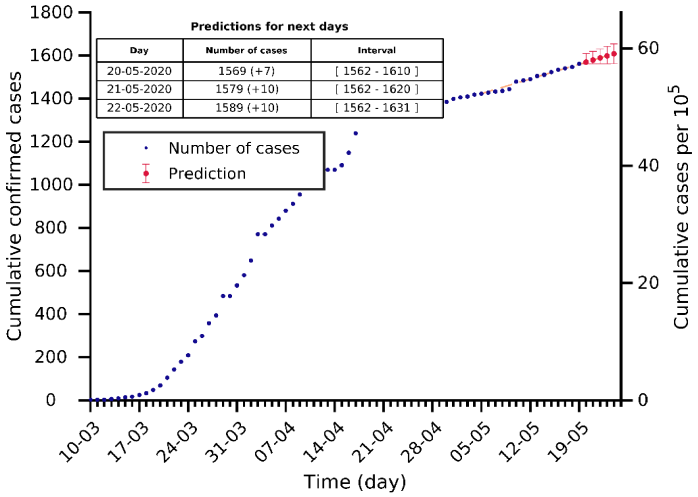


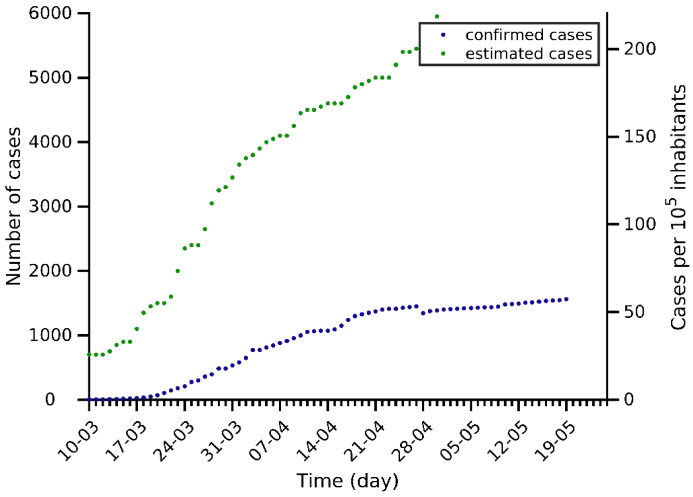


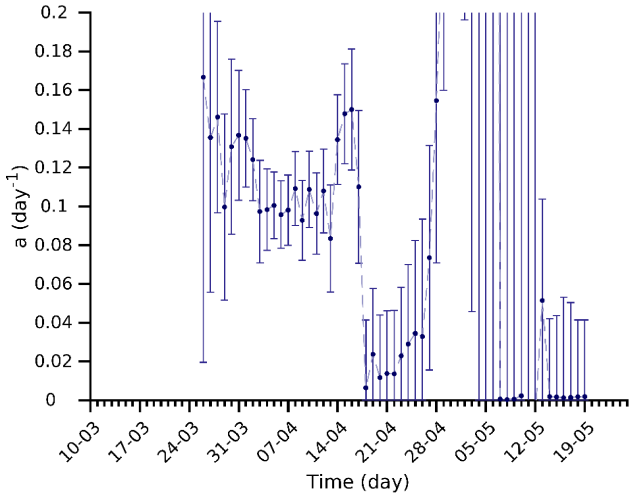


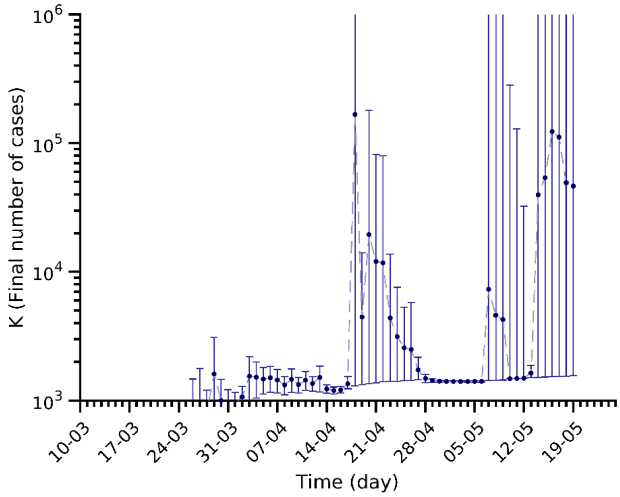


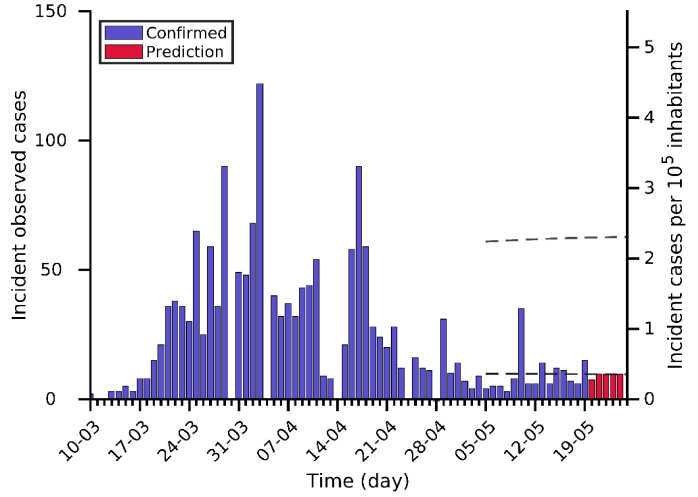
Lithuania 19-05-2020. Population: 2.7M. Current cumulated incidence: $57/10^5$

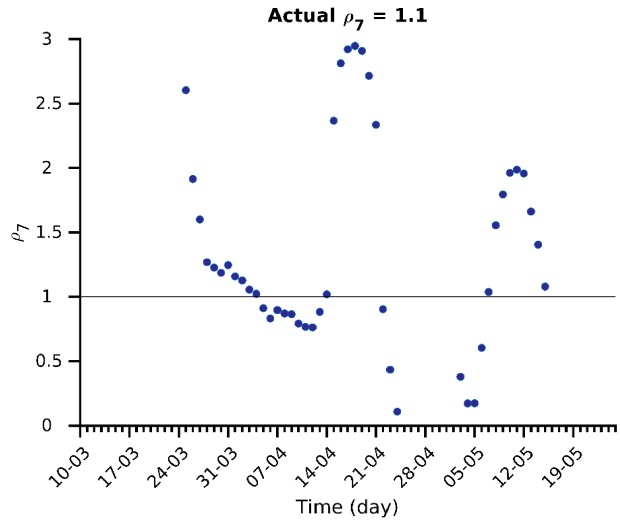


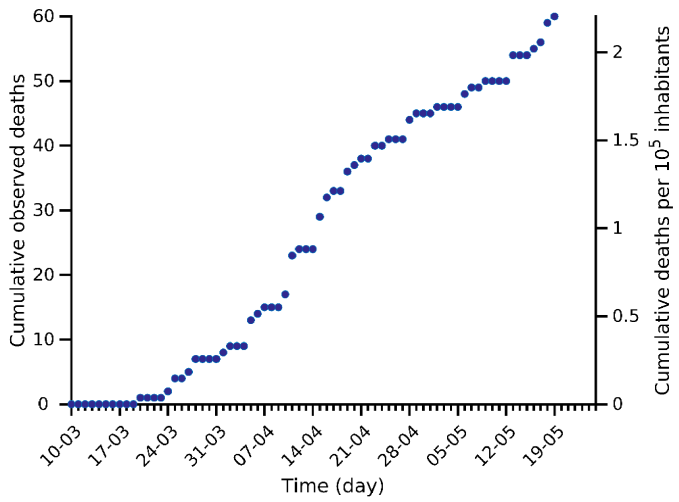


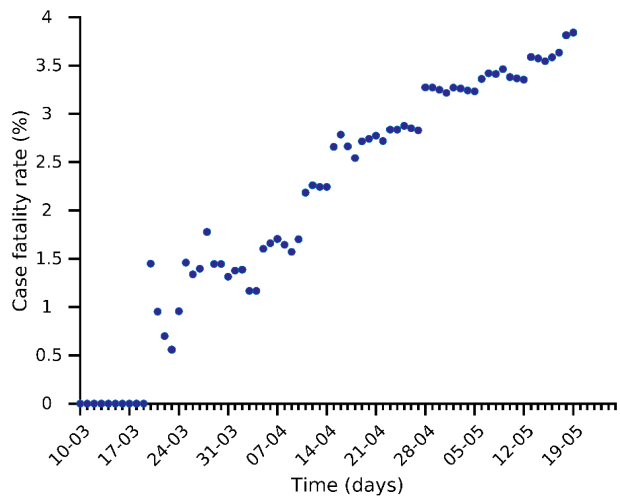




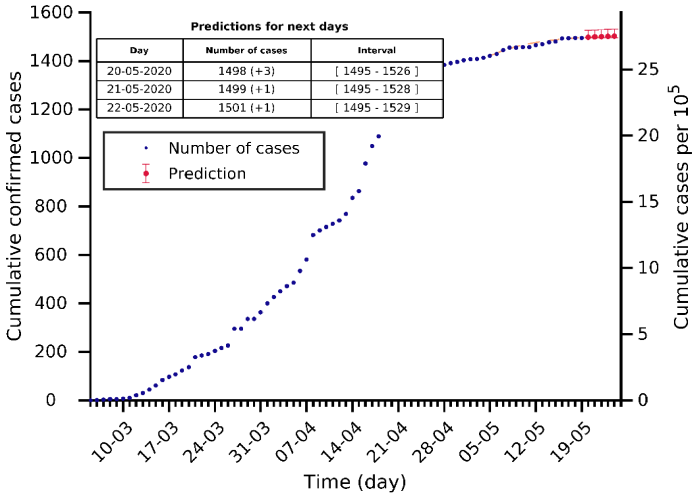


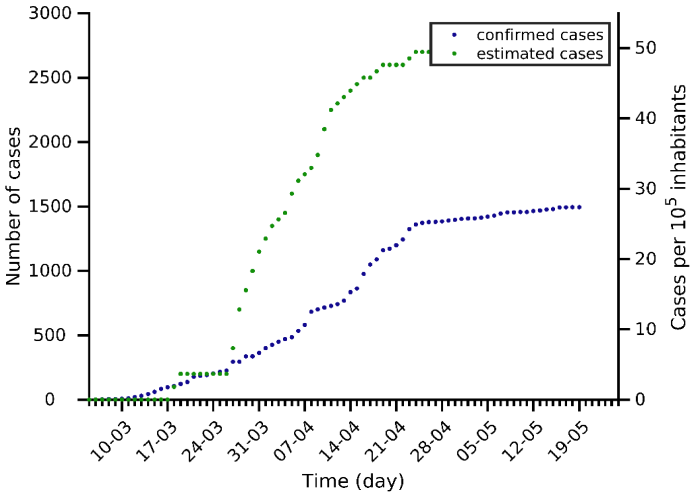


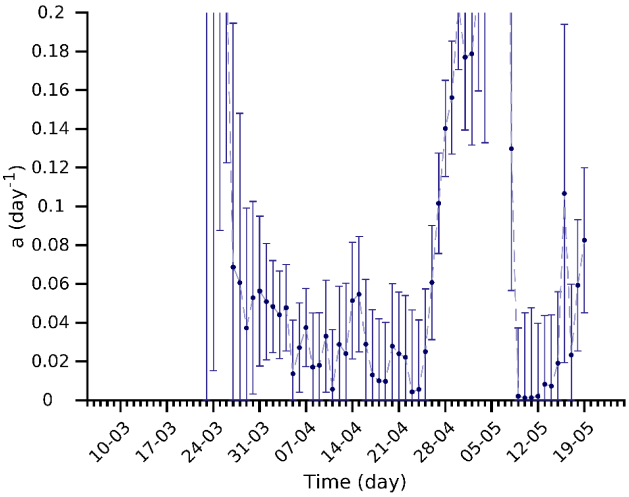


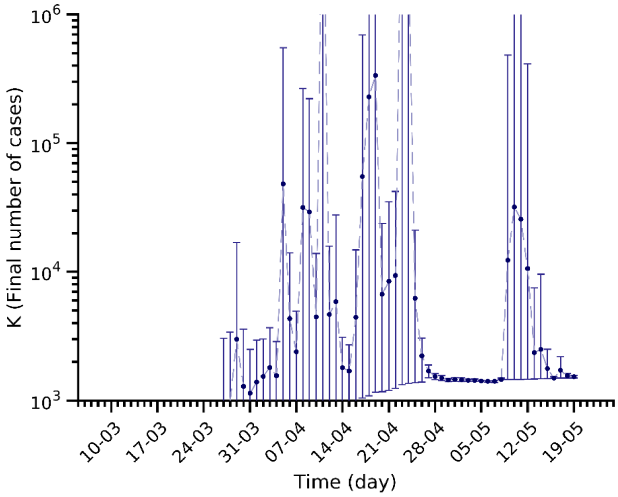


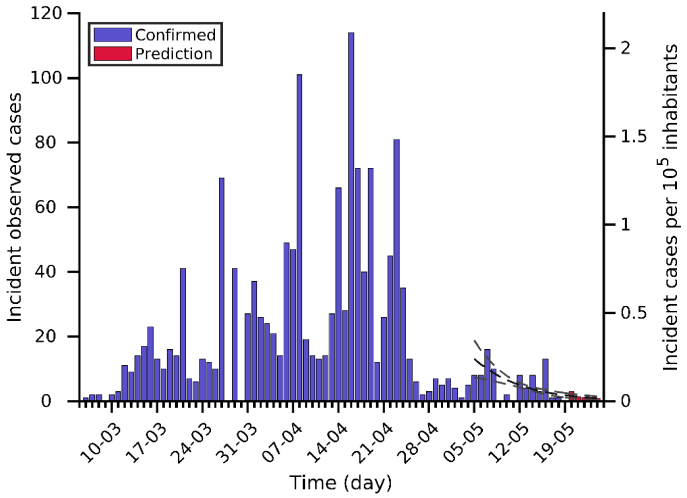
Slovakia 19-05-2020. Population: 5.5M. Current cumulated incidence: 27/10^5

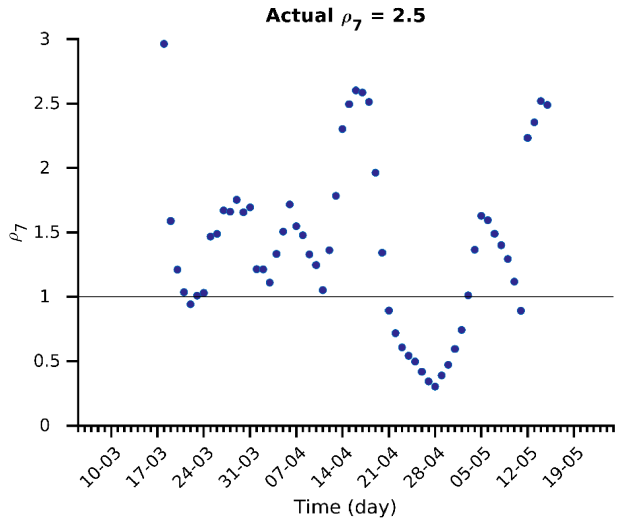


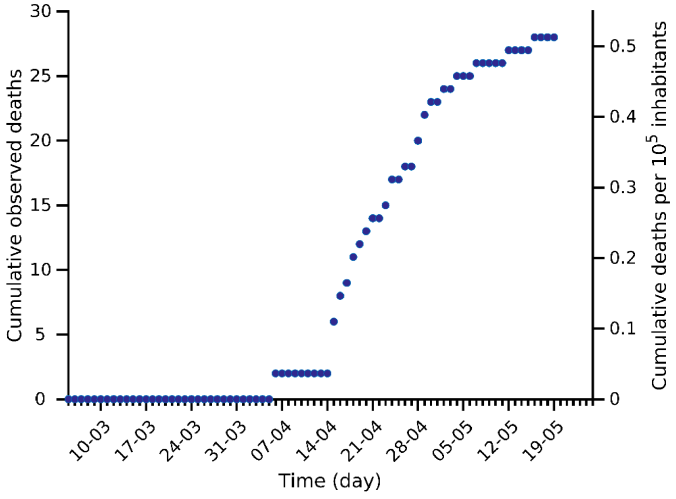


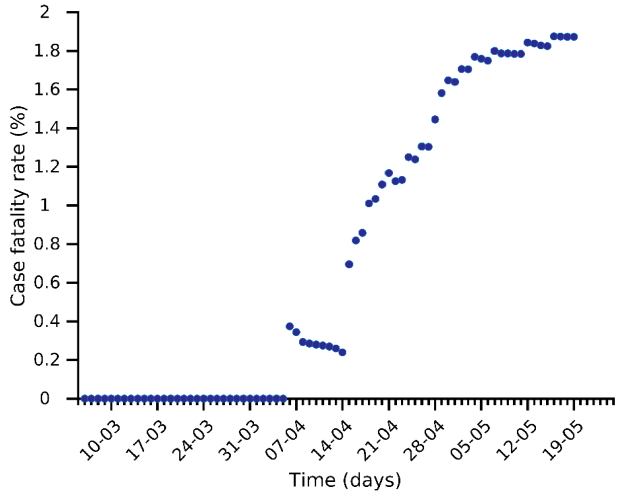




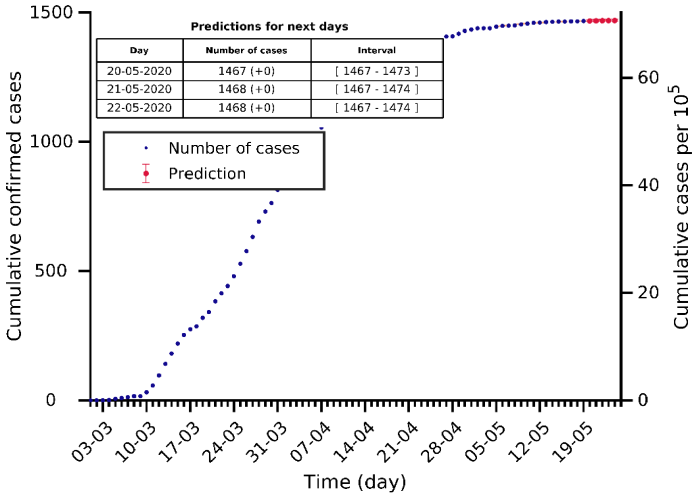


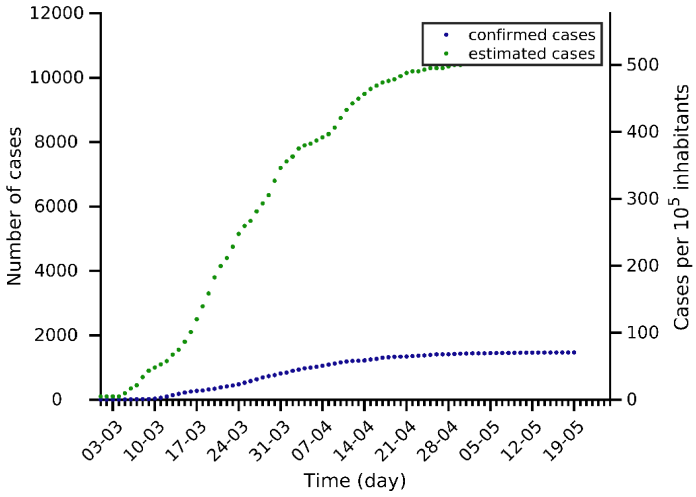


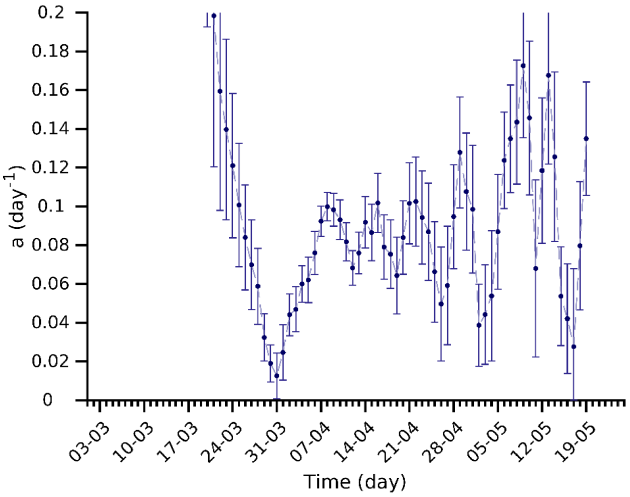


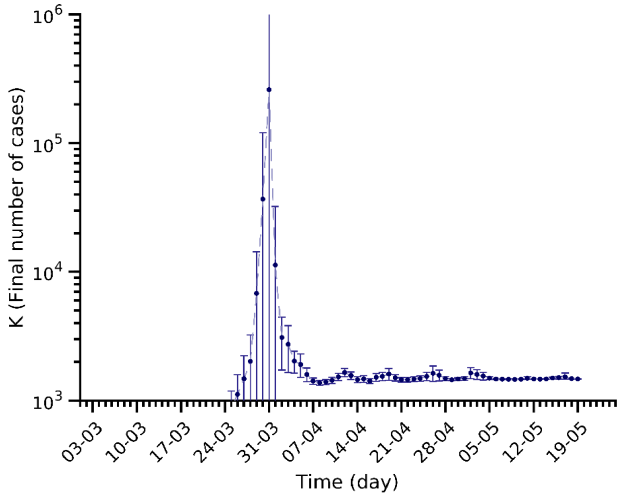


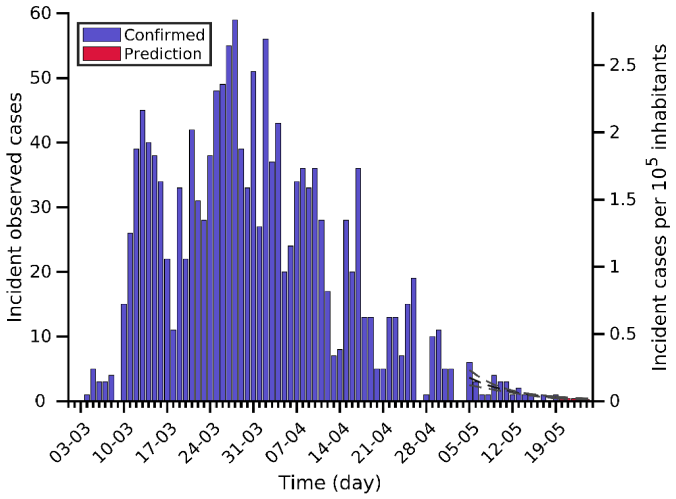
Slovenia 19-05-2020. Population: 2.1M. Current cumulated incidence: 71/10^5

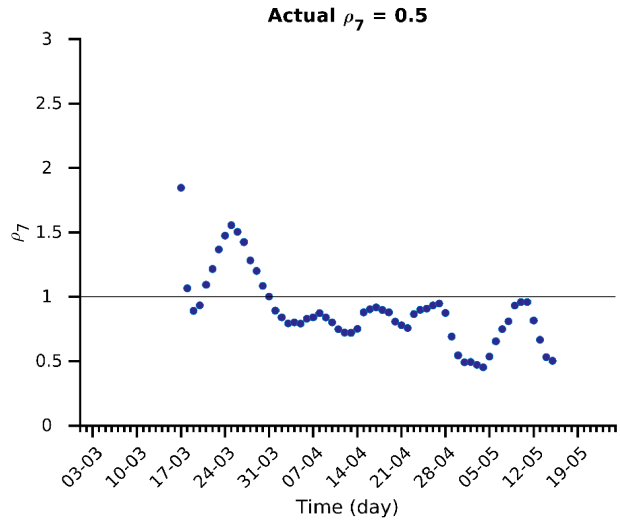


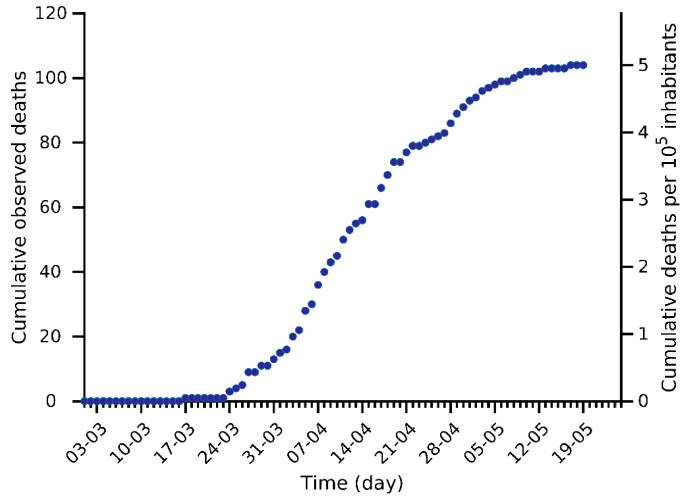


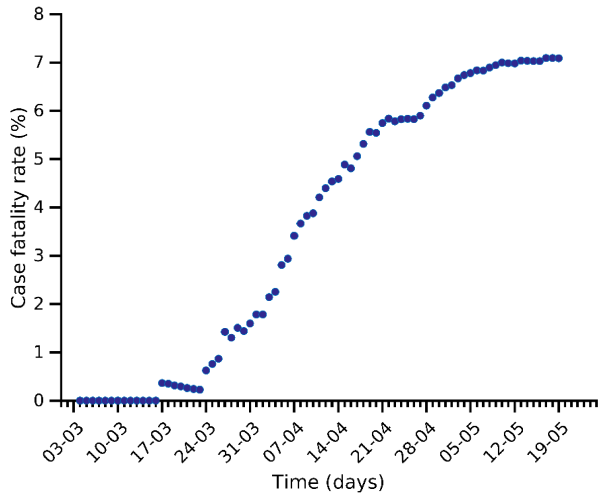




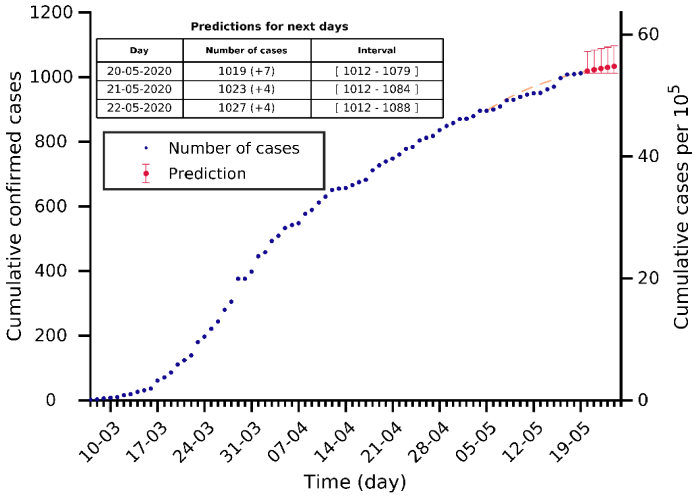


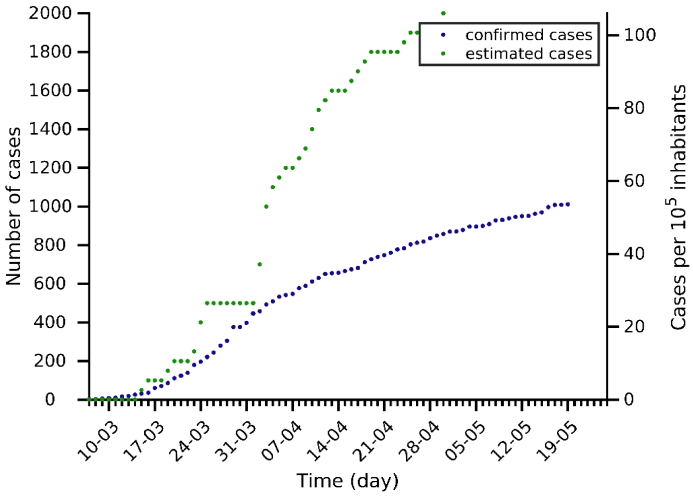


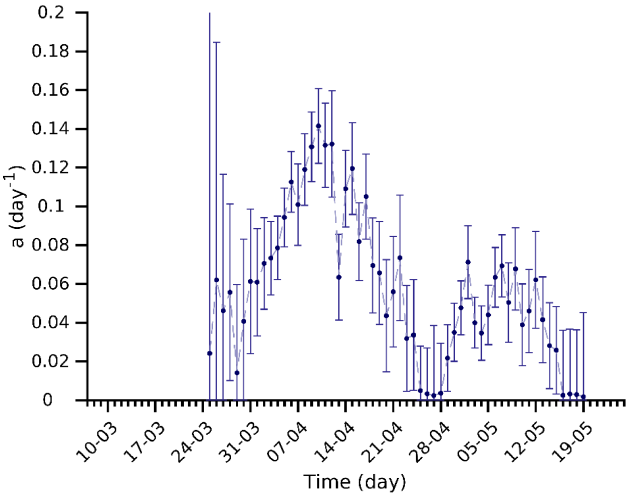


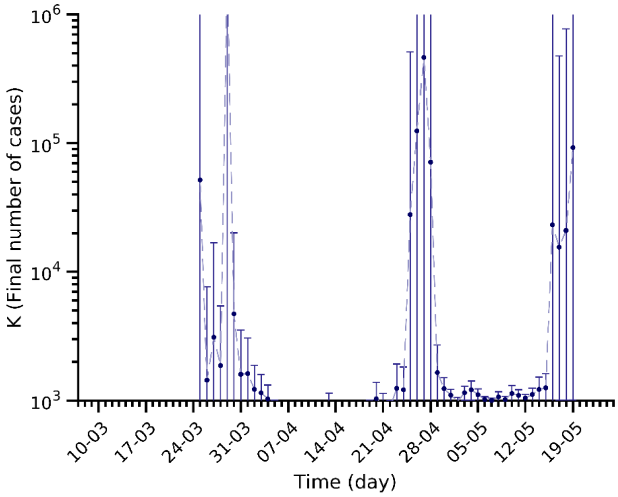


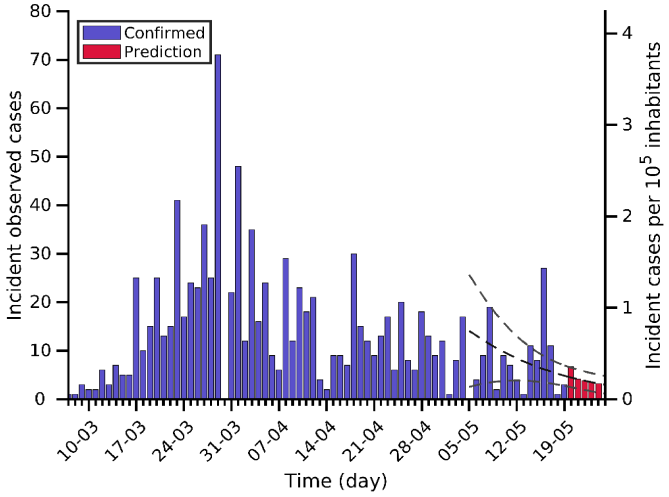
Latvia 19-05-2020. Population: 1.9M. Current cumulated incidence: $54/10^5$

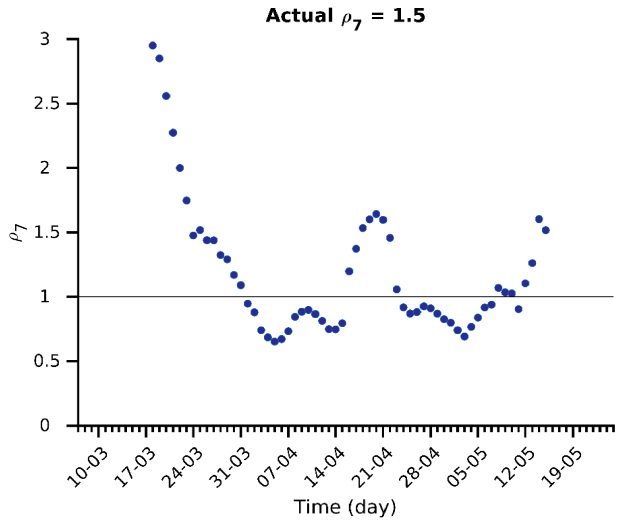


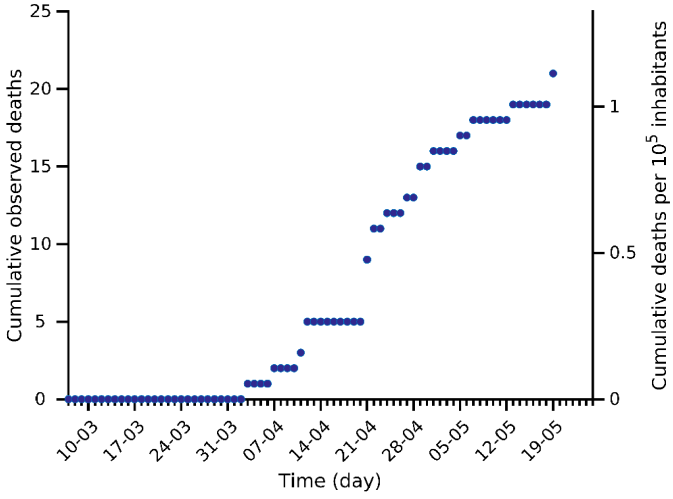


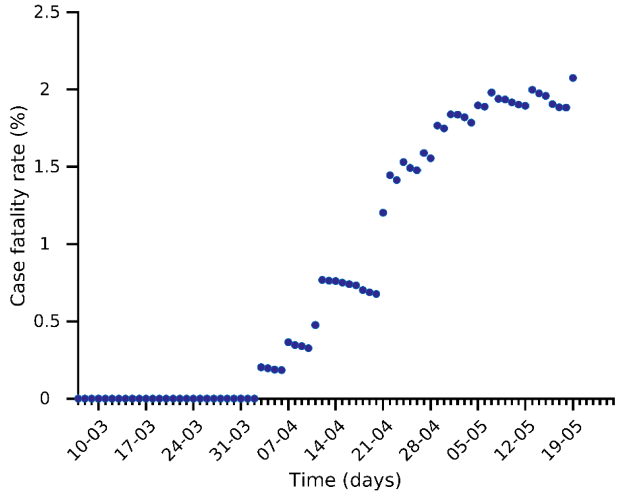




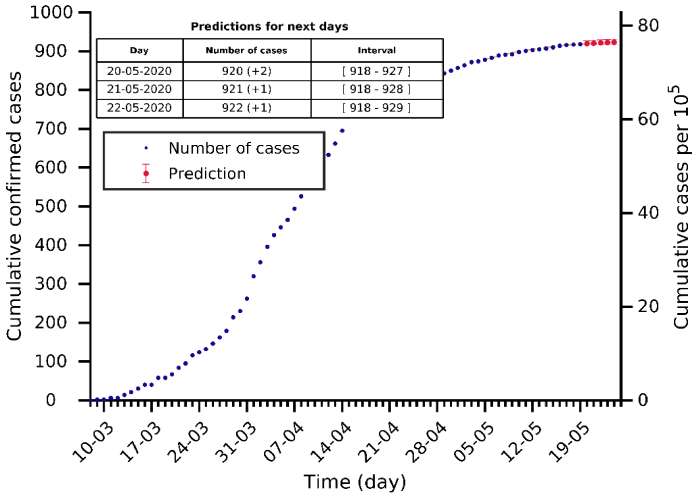


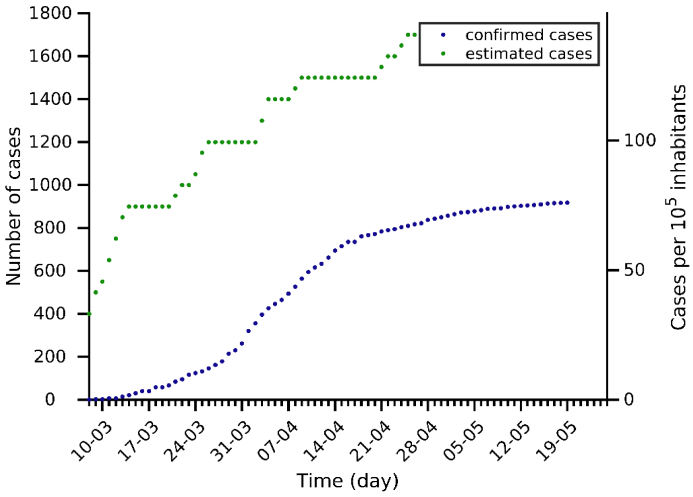


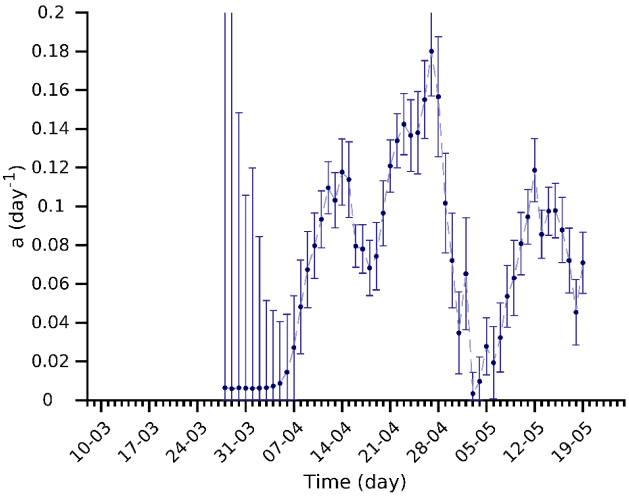


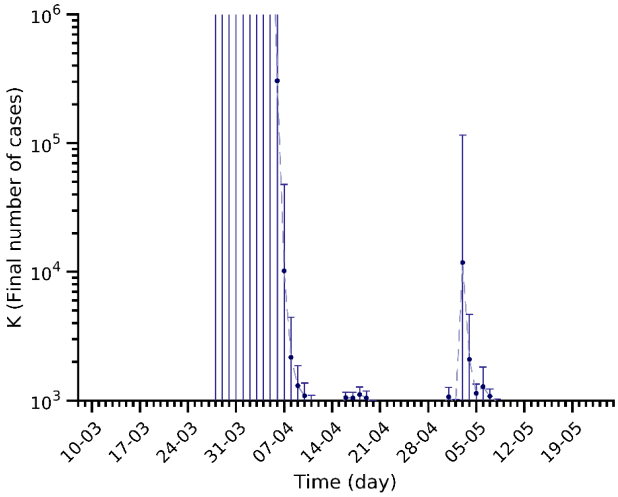


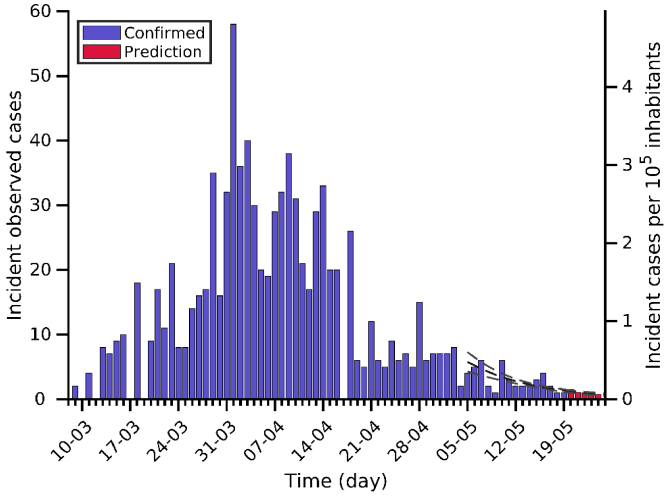
Cyprus 19-05-2020. Population: 1.2M. Current cumulated incidence: 76/10^5

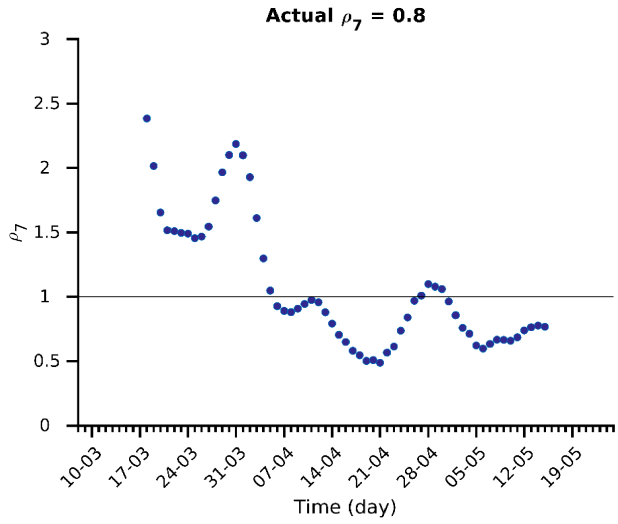


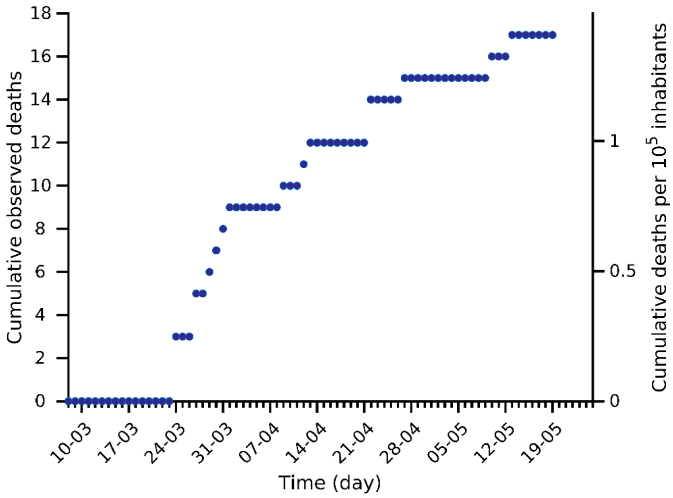


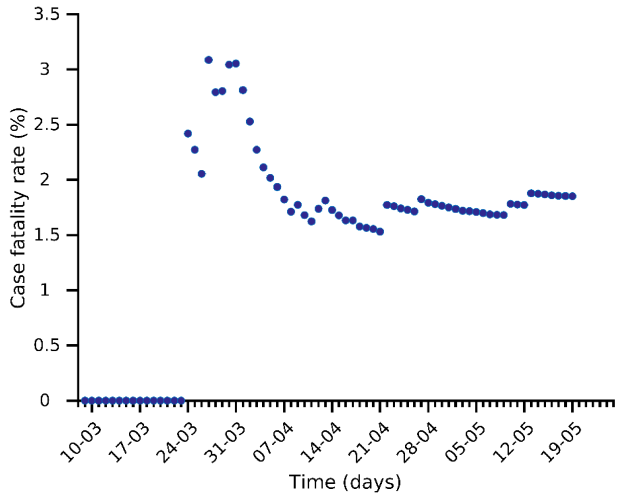




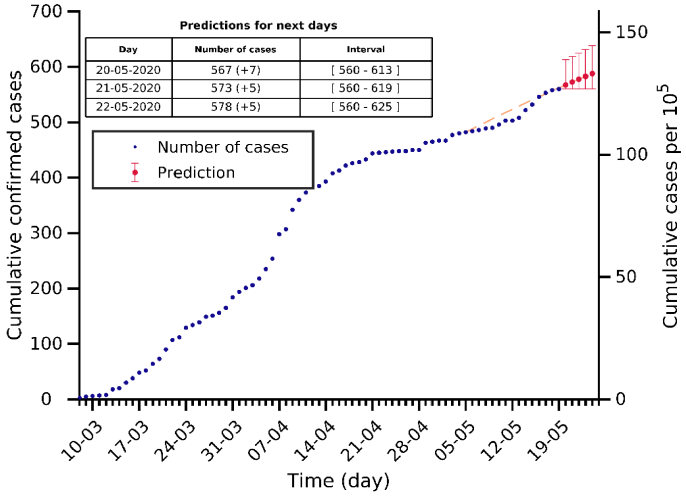


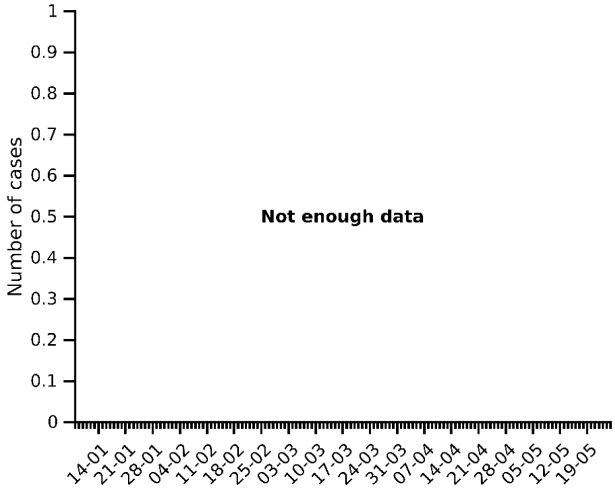


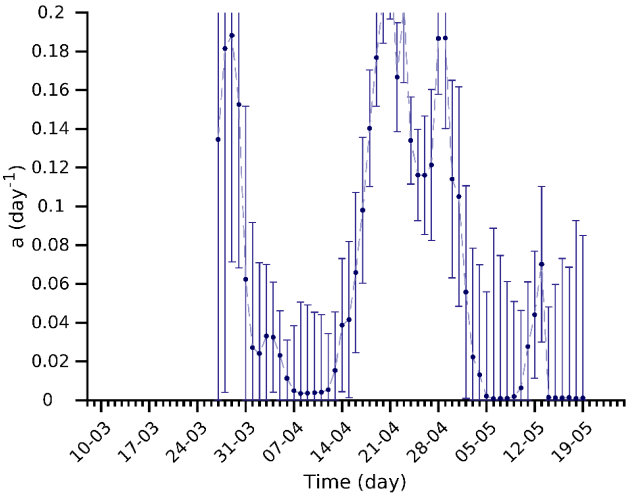


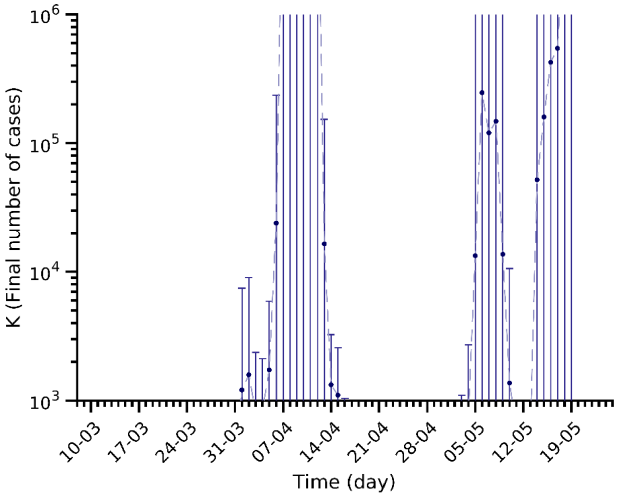


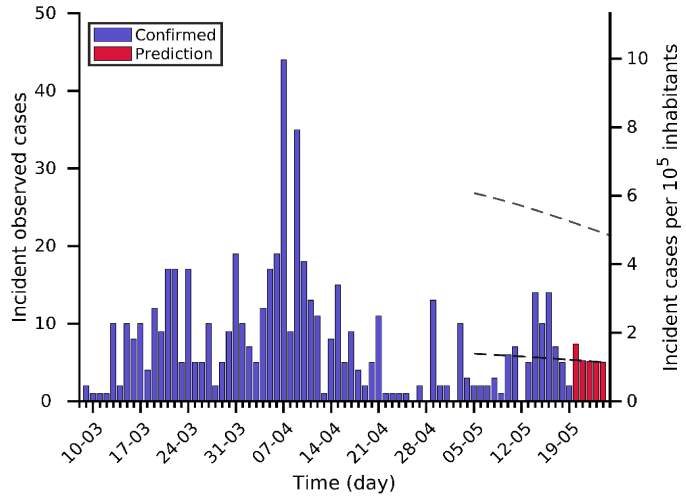
Malta 19-05-2020. Population: 0.4M. Current cumulated incidence: $127/10^5$

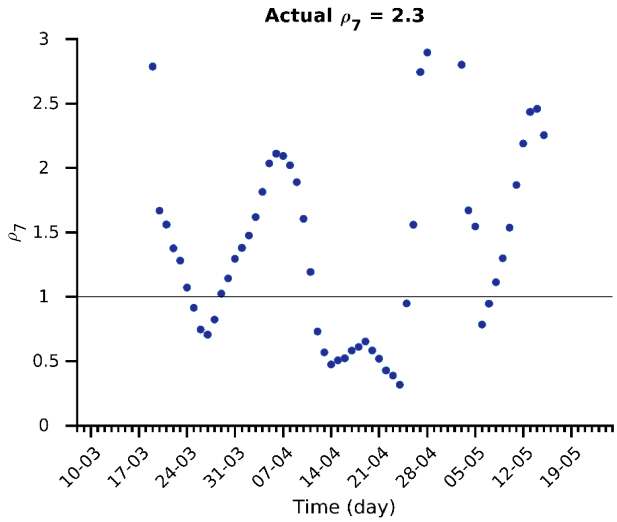


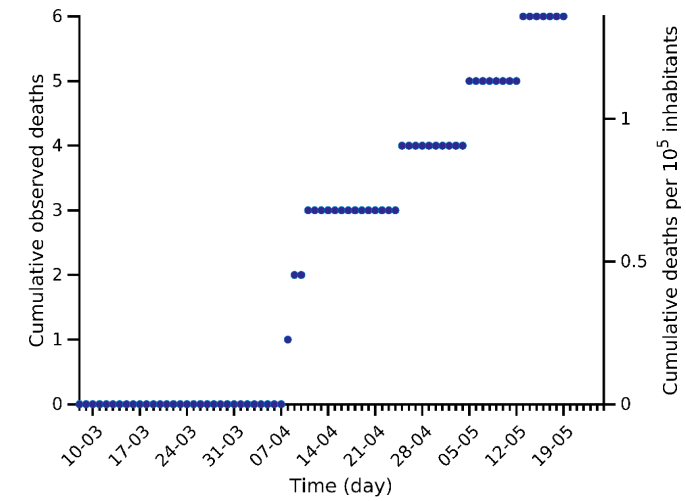


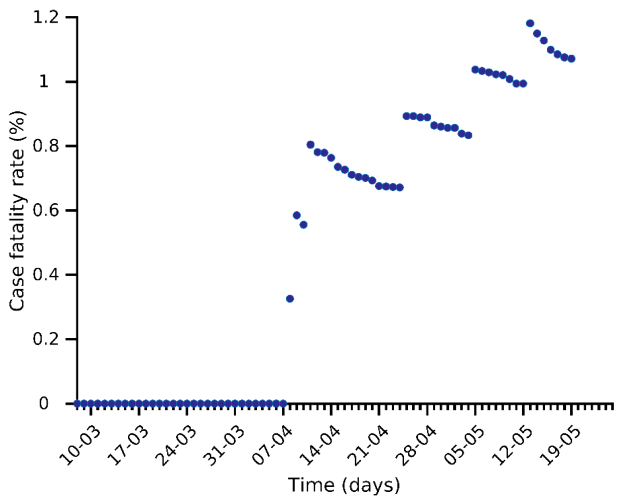








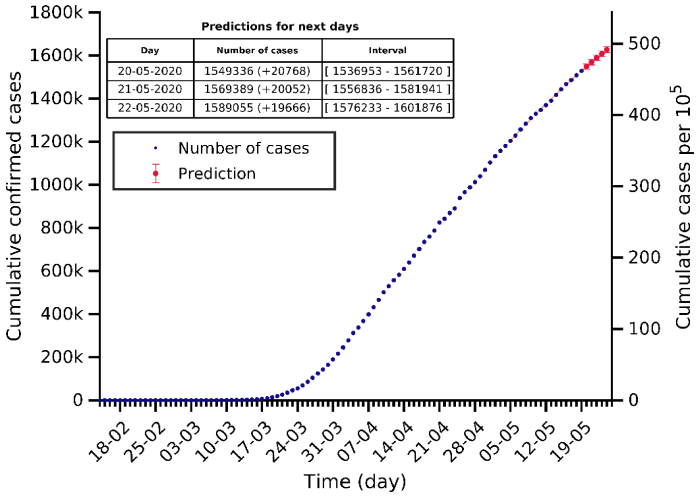


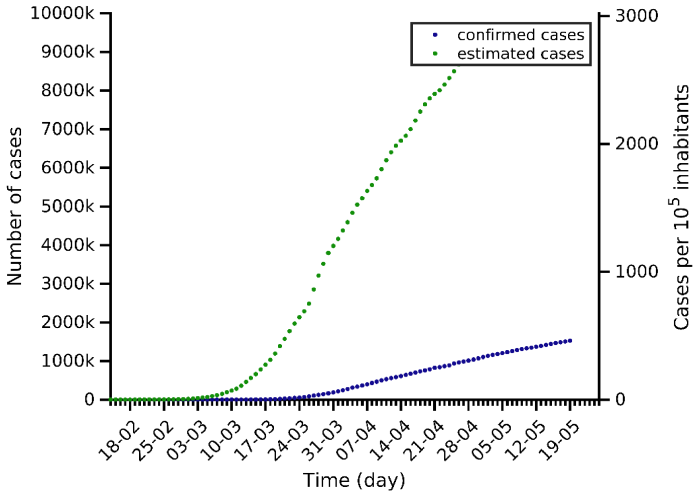


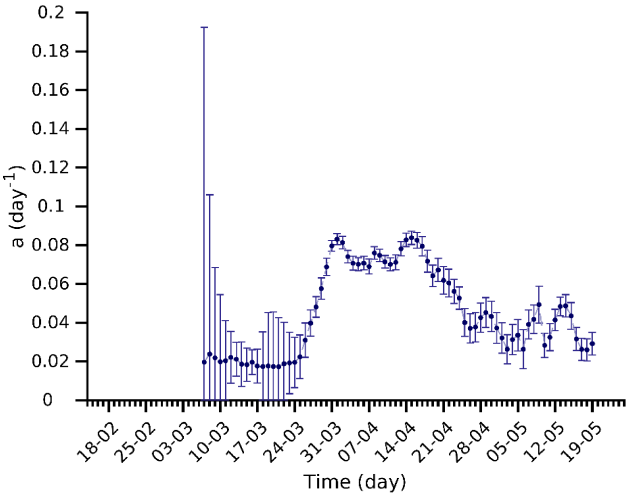
(2) Analysis and prediction of COVID-19 for other countries

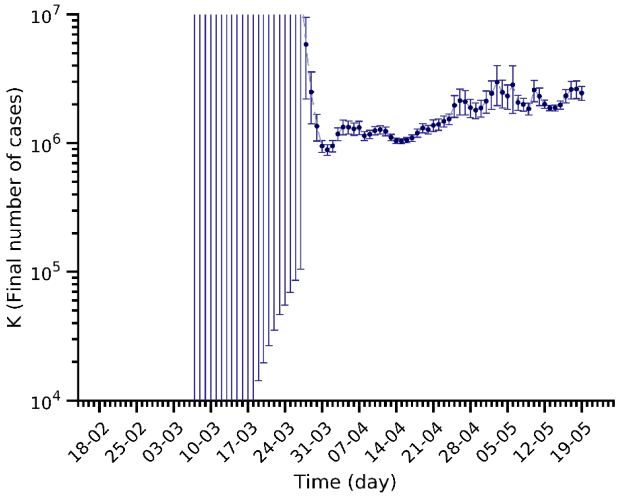
Data obtained from https://www.ecdc.europa.eu/en/geographical-distribution-2019-ncov-cases

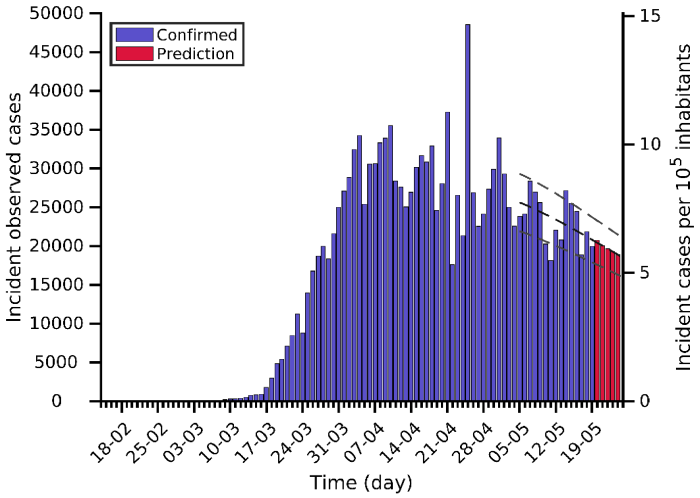
USA 19-05-2020. Population: 331.0M. Current cumulated incidence: 462/10^5

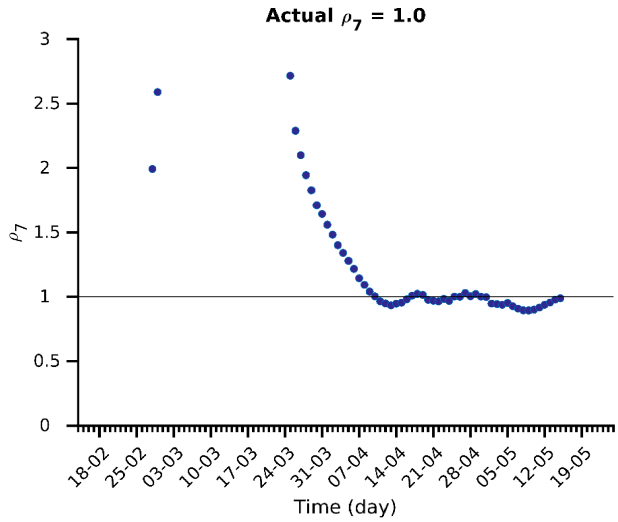


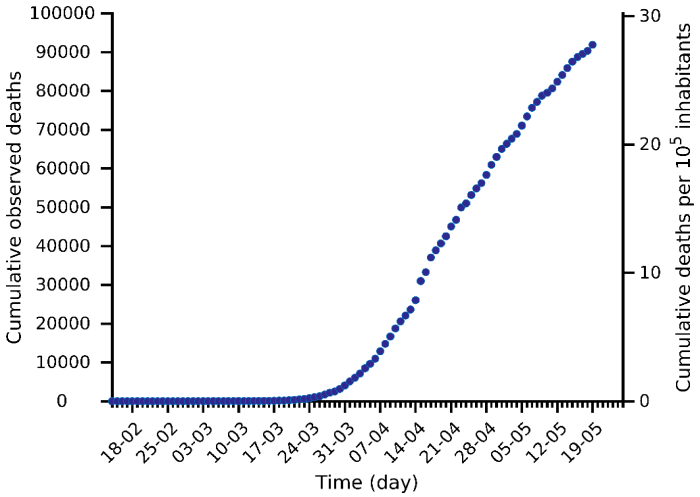


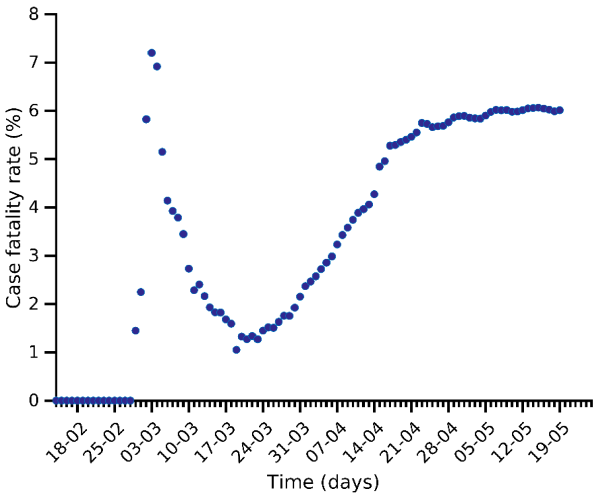




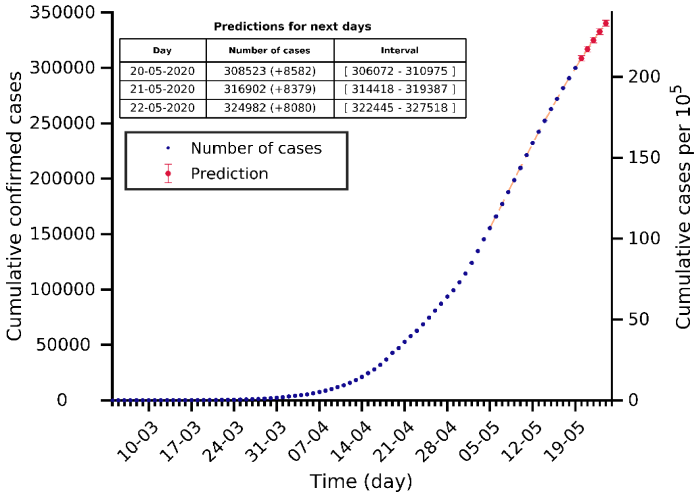


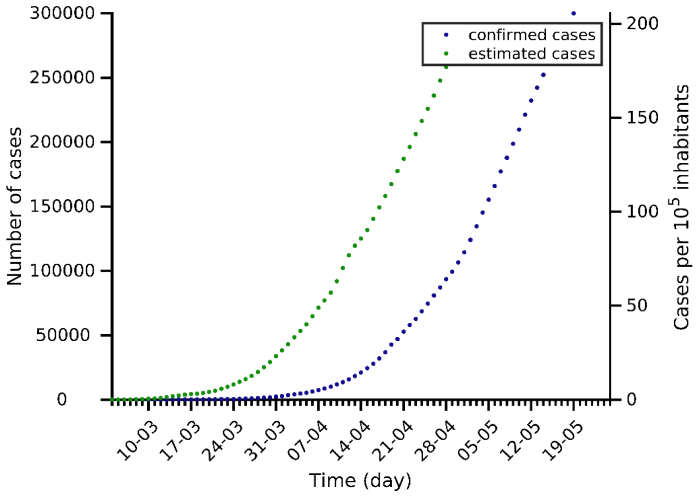


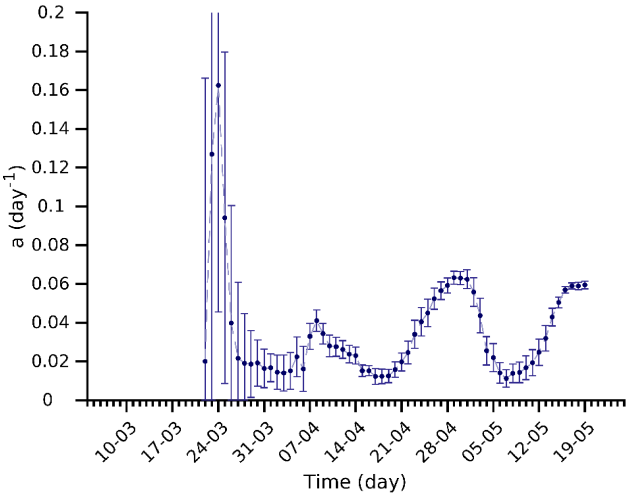


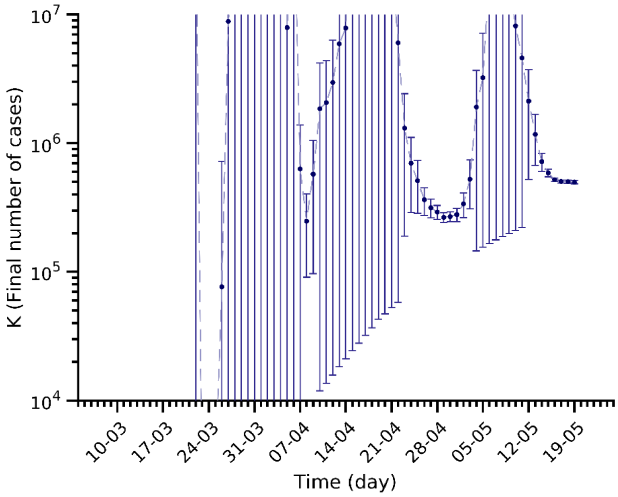


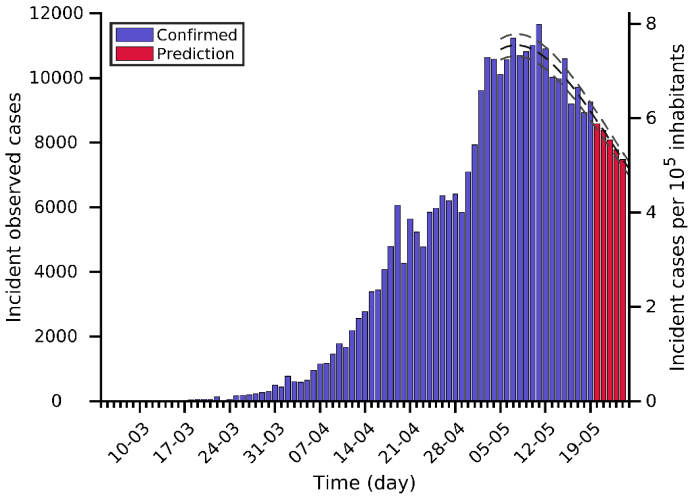
Russia 19-05-2020. Population: 145.9M. Current cumulated incidence: 206/10^5

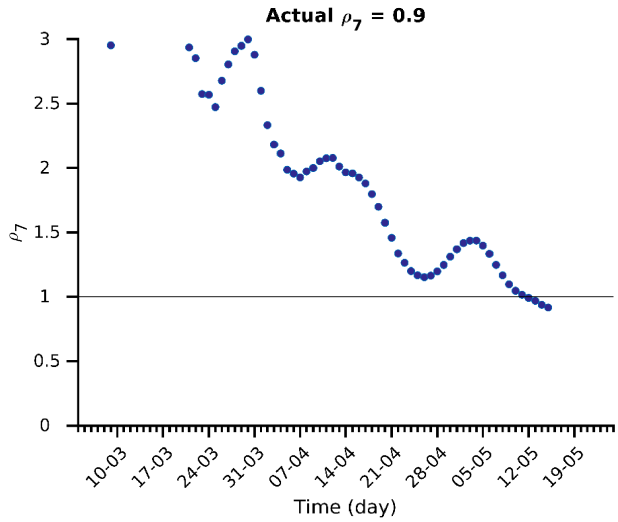


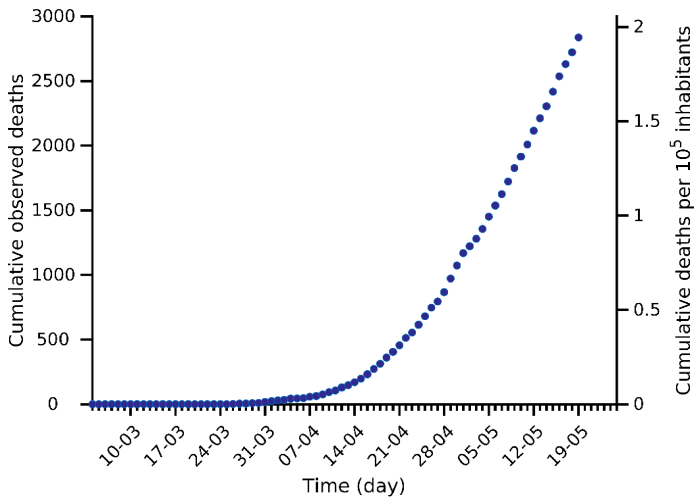


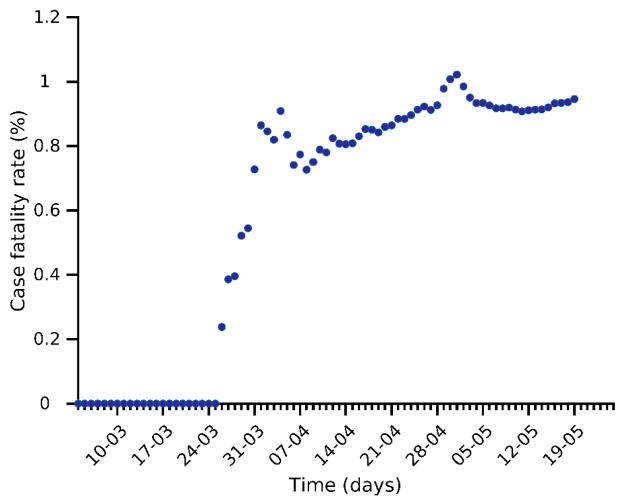




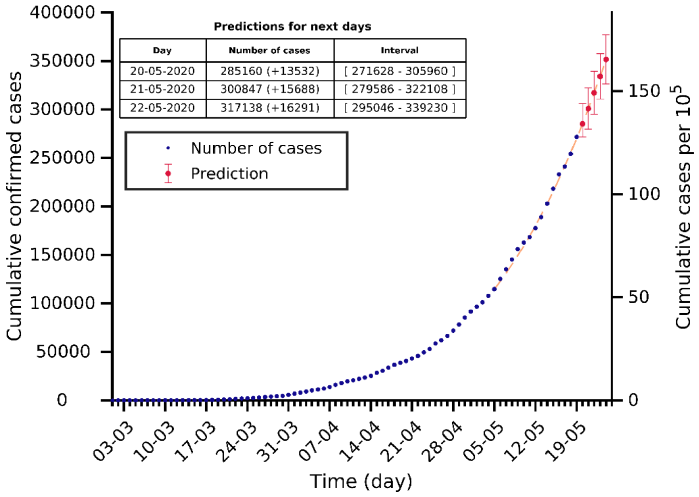


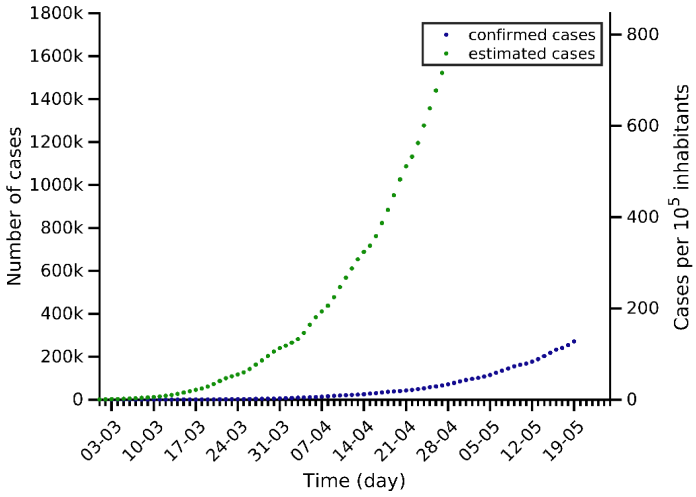


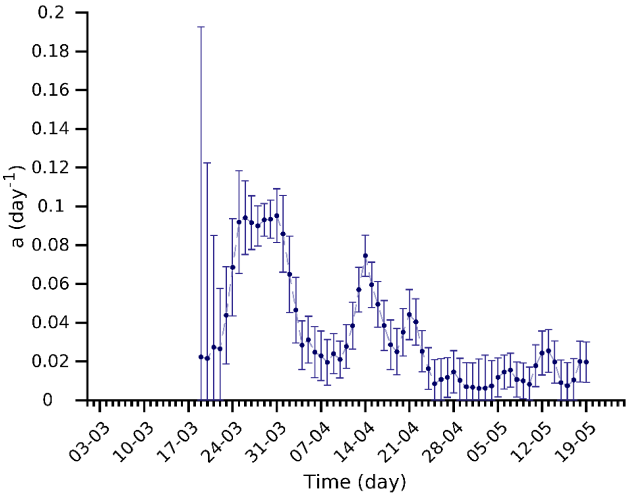


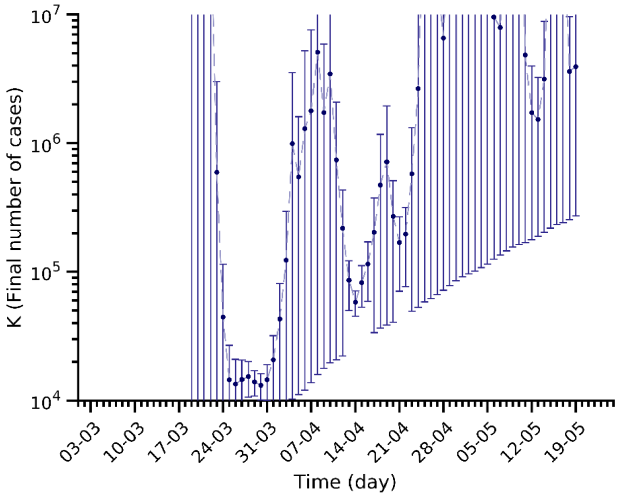


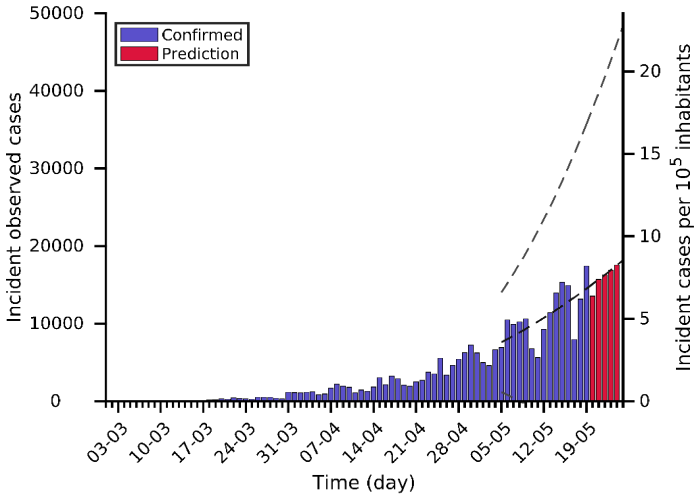
Brazil 19-05-2020. Population: 212.6M. Current cumulated incidence: 128/10^5

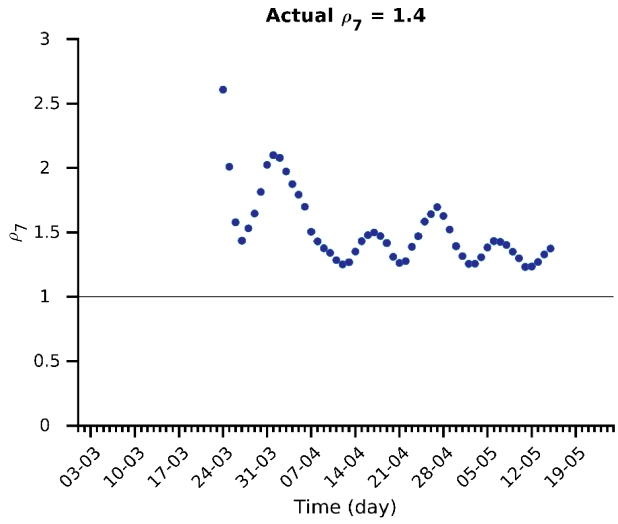


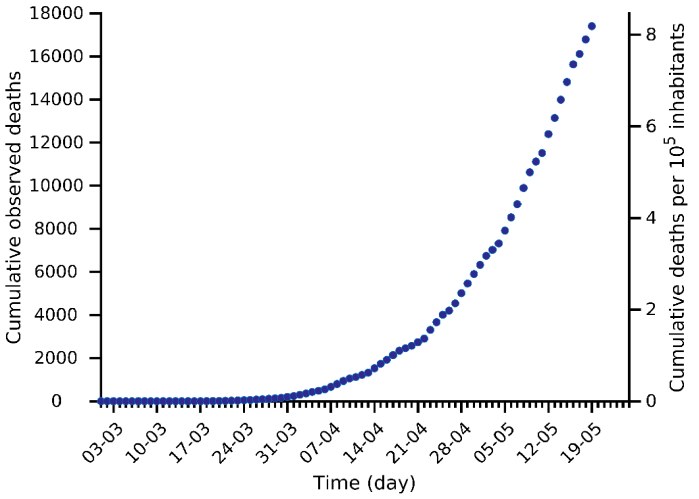


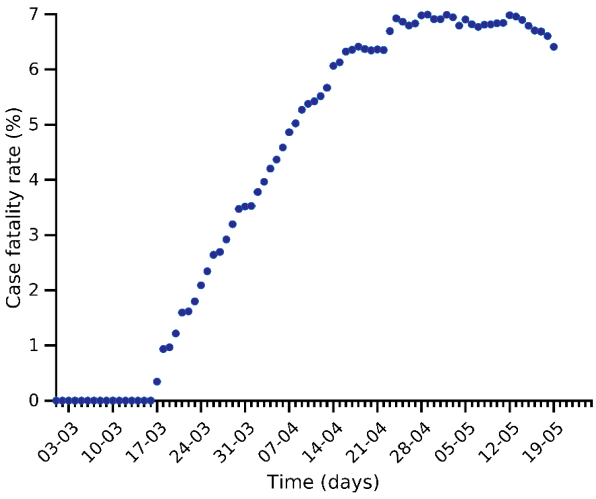




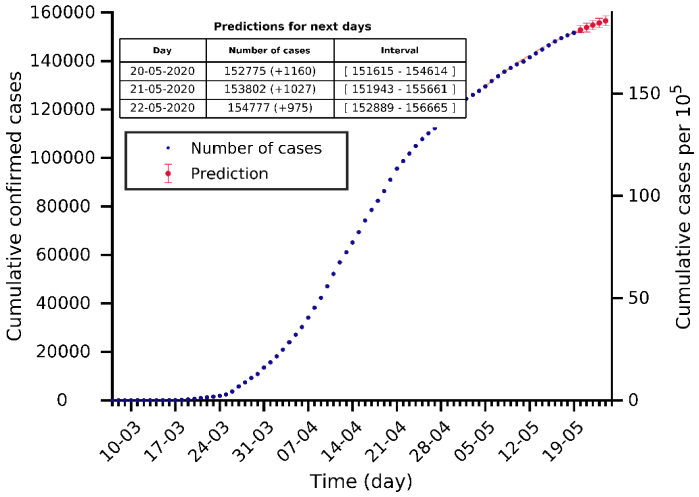


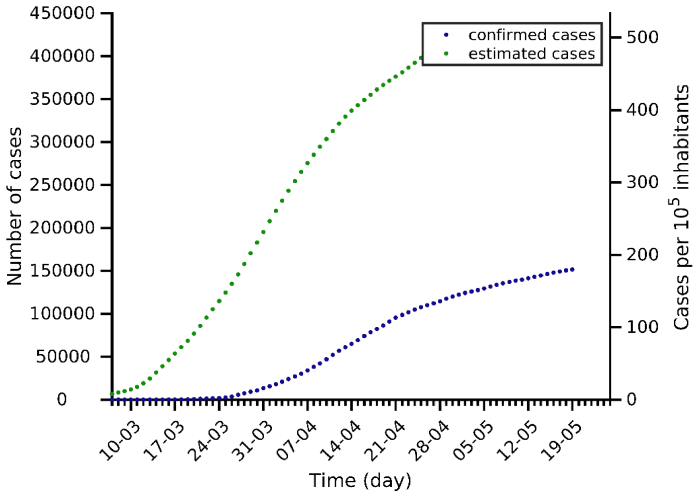


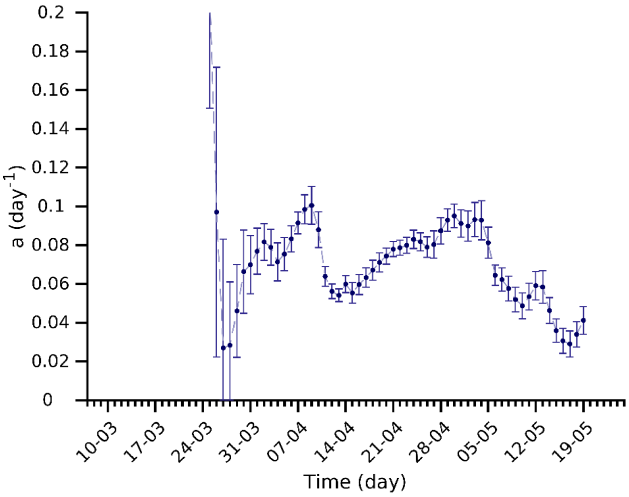


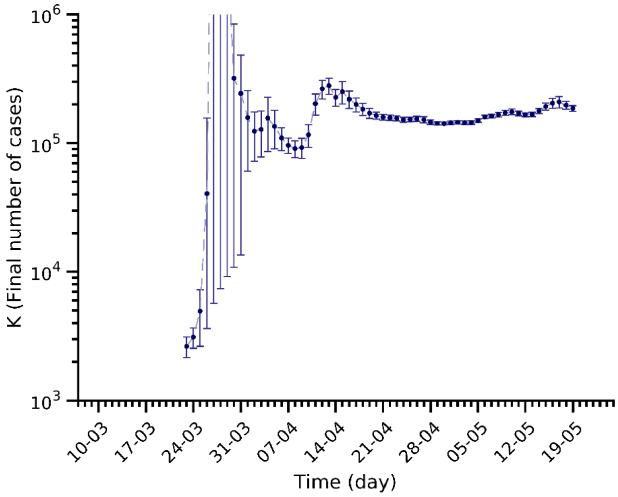


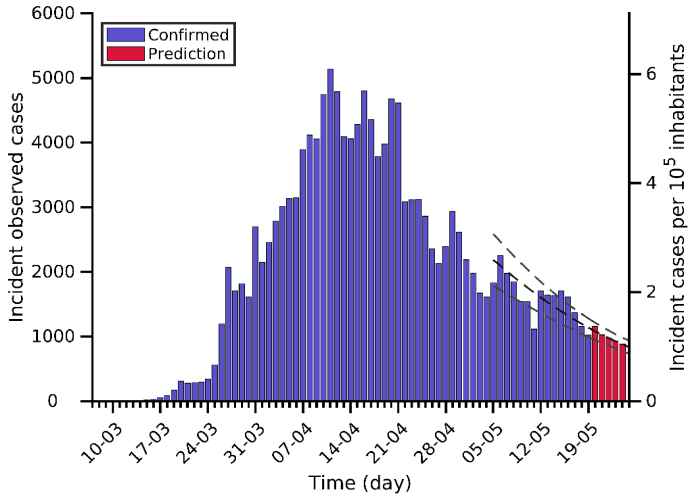
Turkey 19-05-2020. Population: 84.3M. Current cumulated incidence: 180/10^5

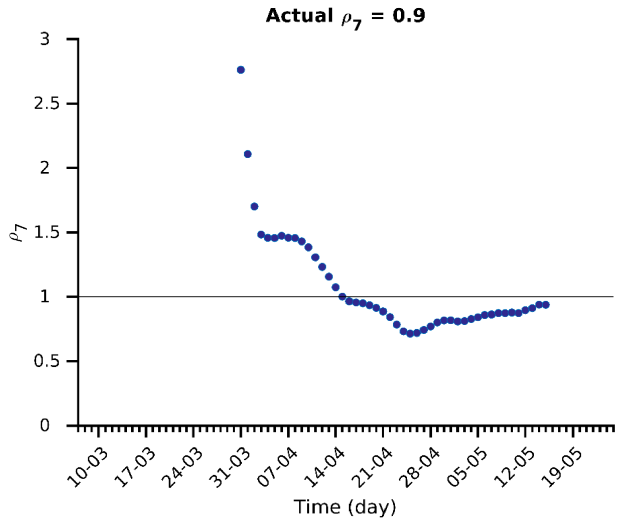


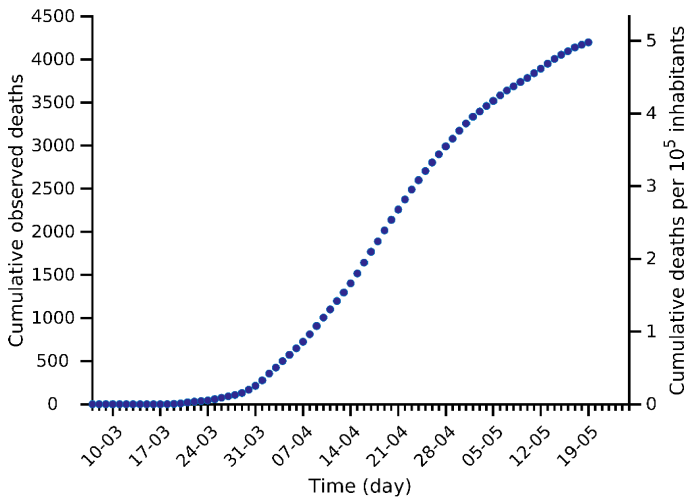


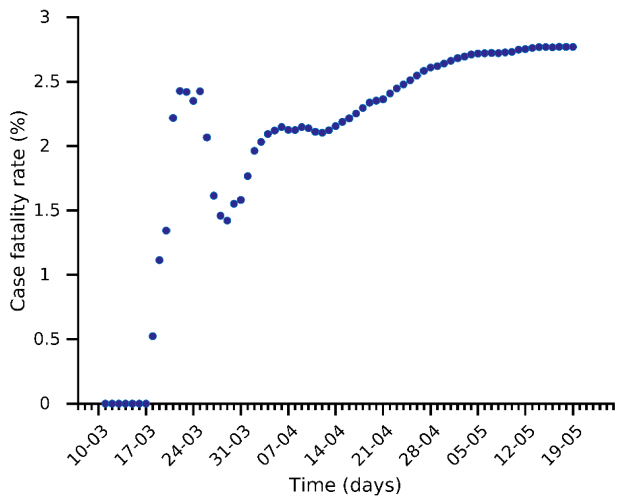




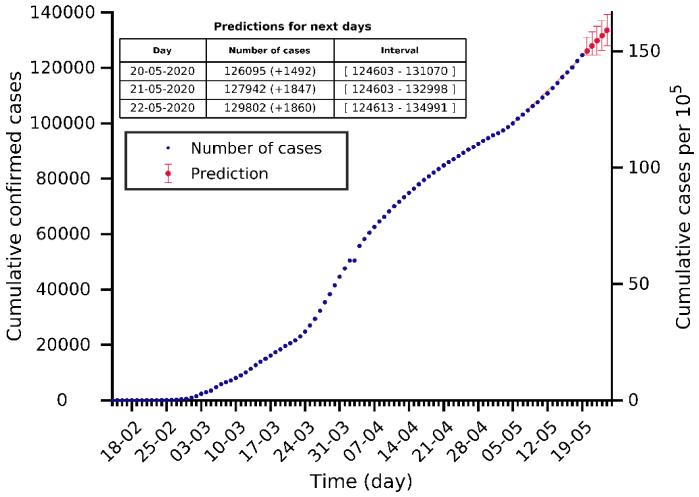


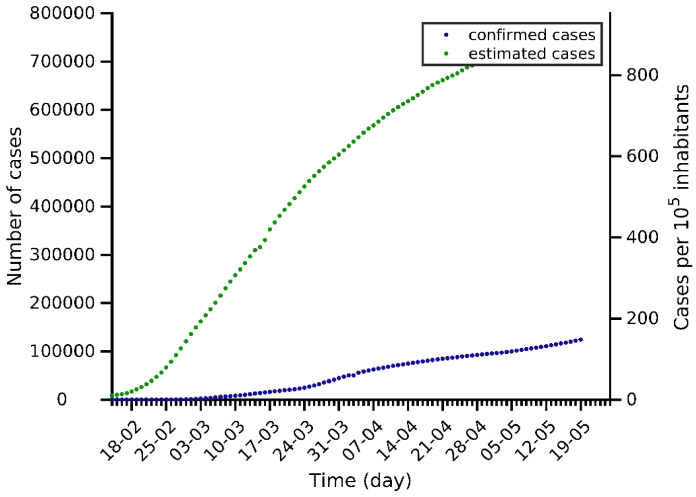


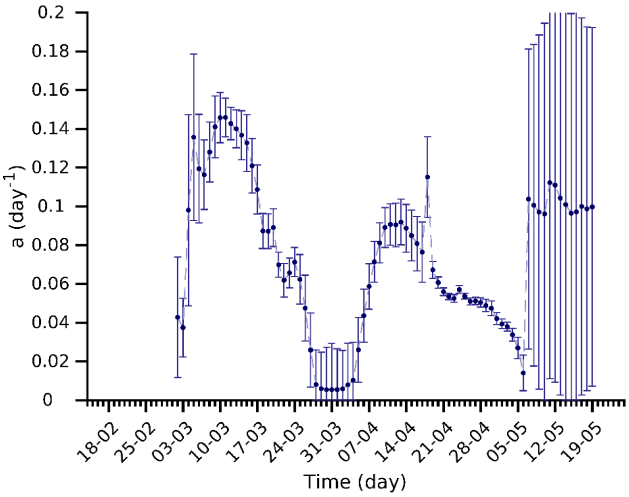


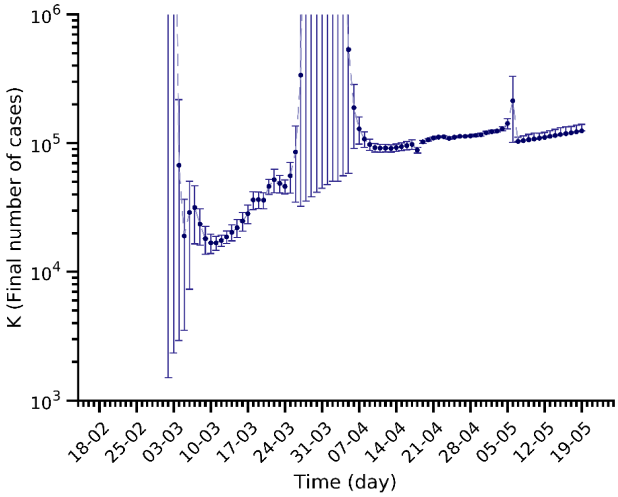


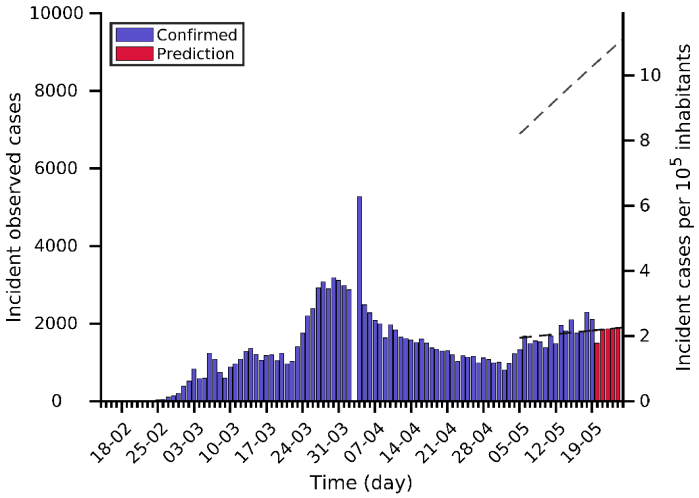
Iran 19-05-2020. Population: 84.0M. Current cumulated incidence: 148/10^5

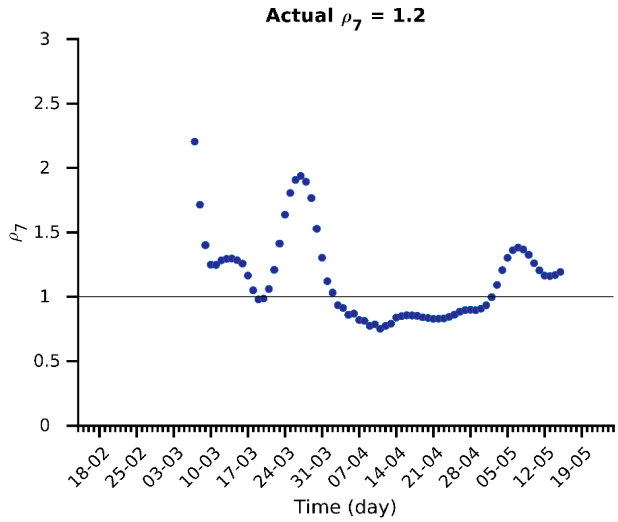


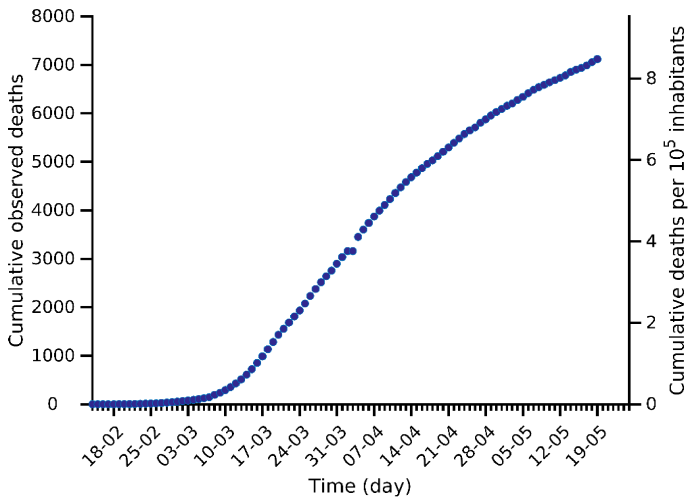


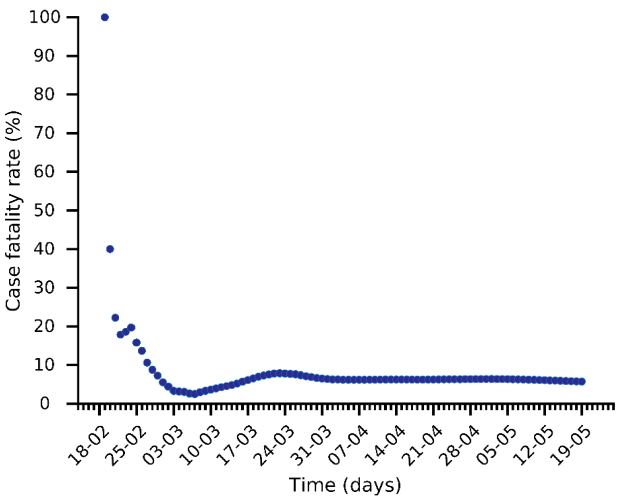




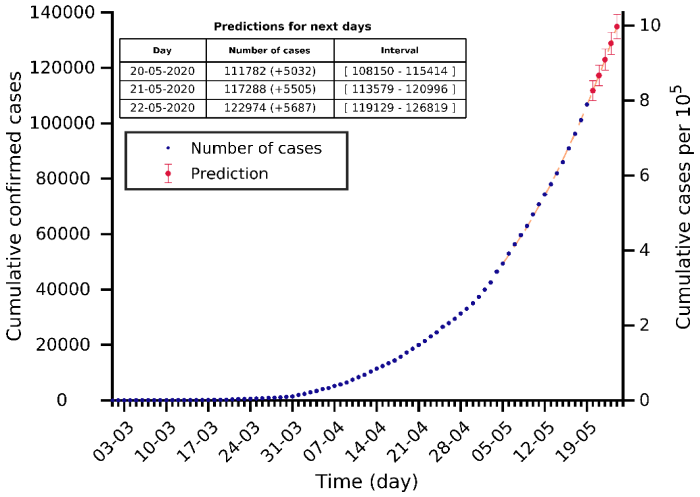


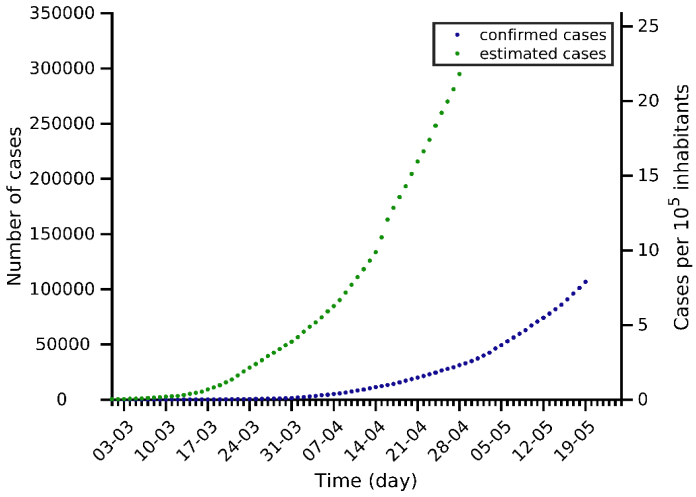


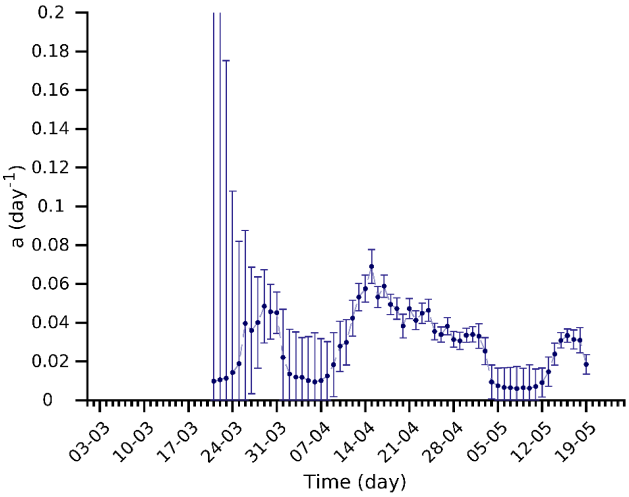


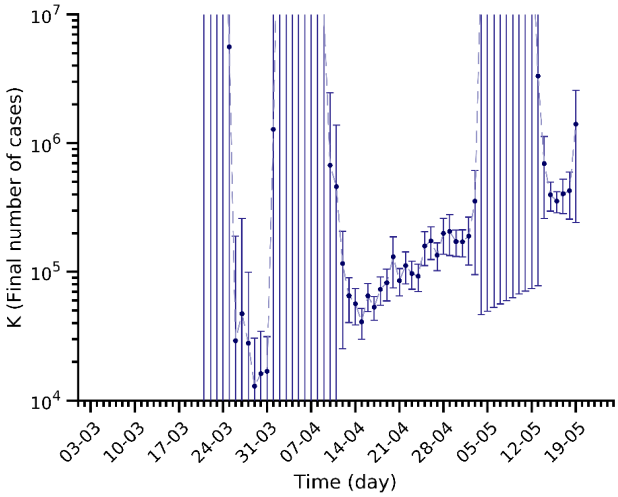


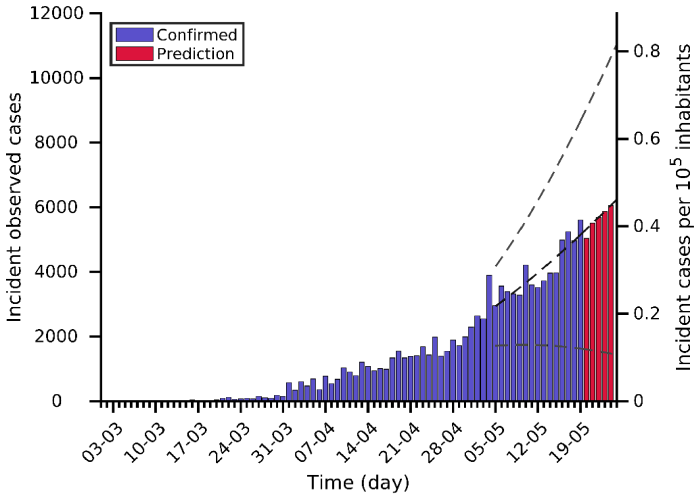
India 19-05-2020. Population: 1353.0M. Current cumulated incidence: $8/10^5$

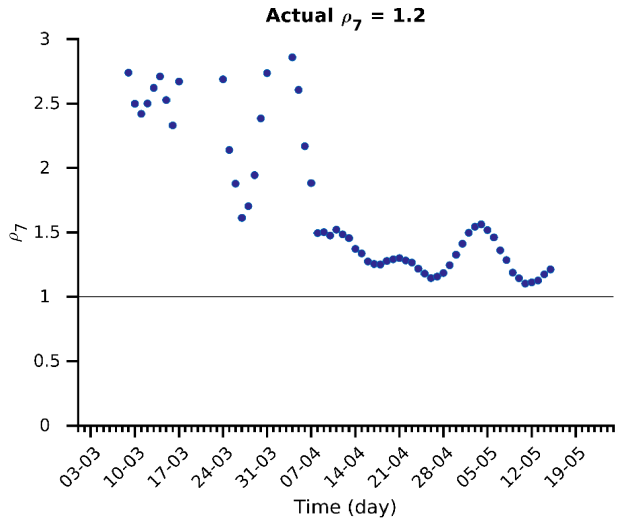


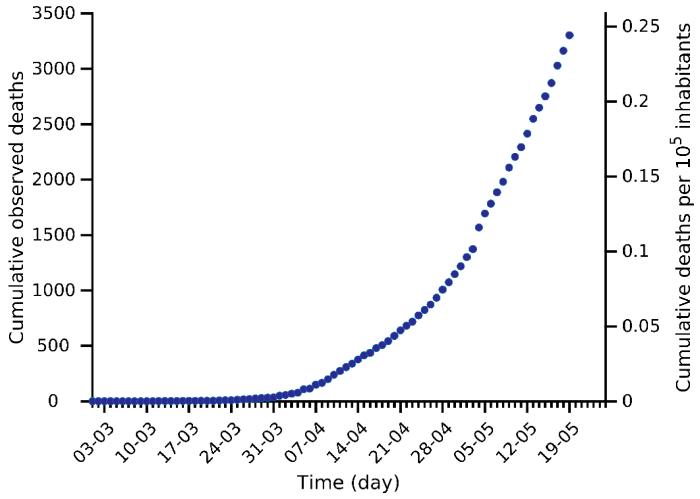


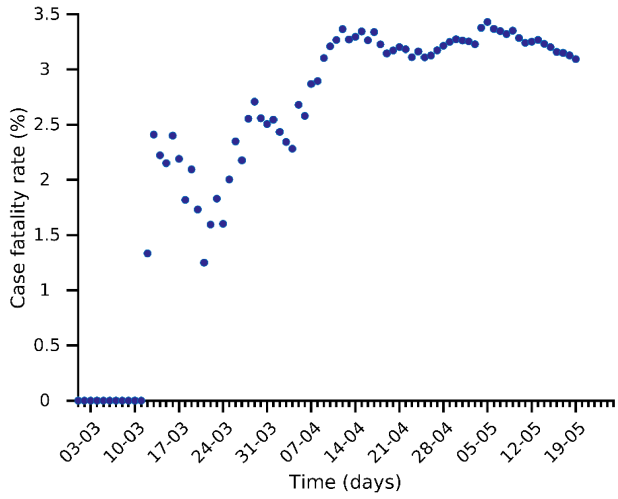




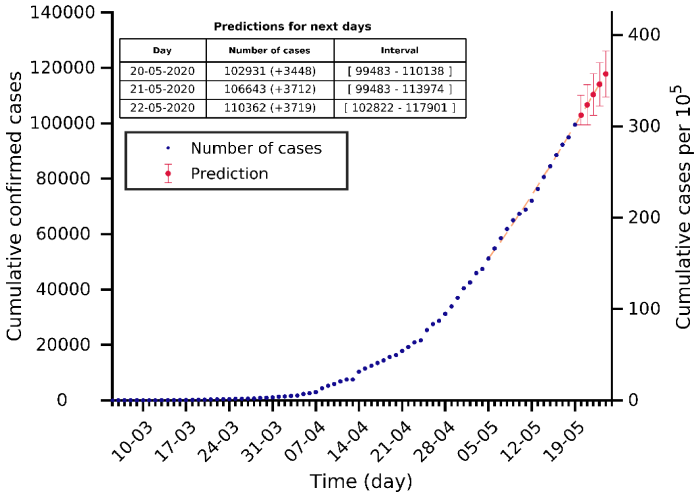


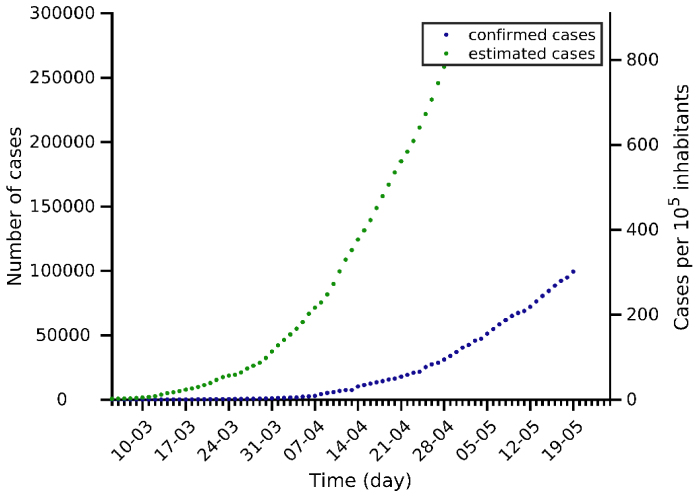


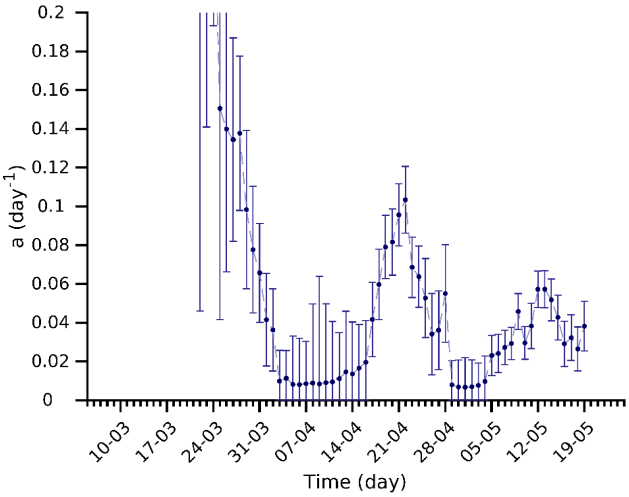


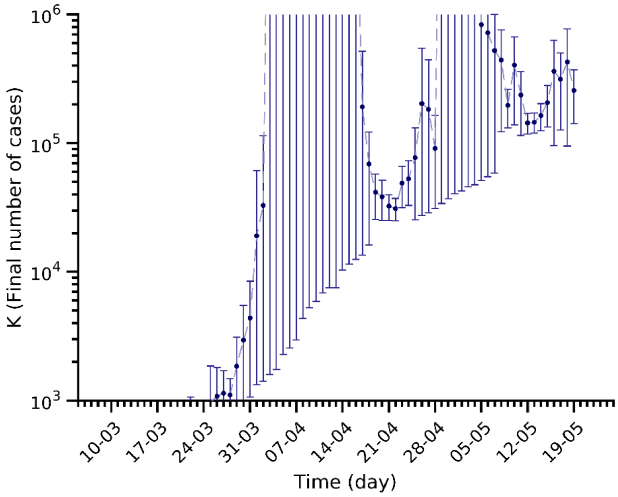


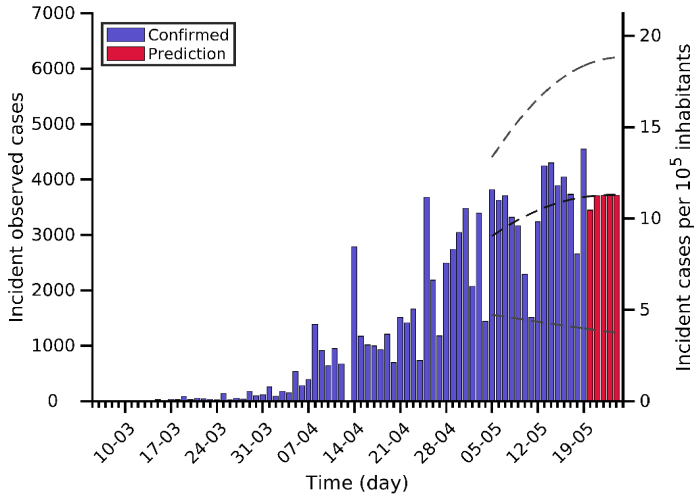
Peru 19-05-2020. Population: 33.0M. Current cumulated incidence: $302/10^5$

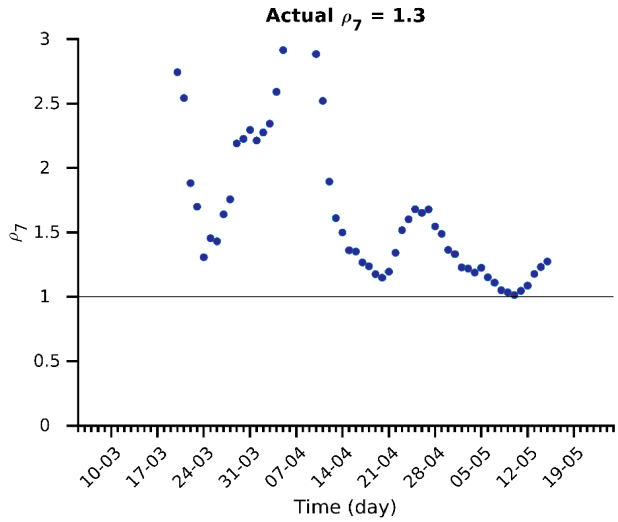


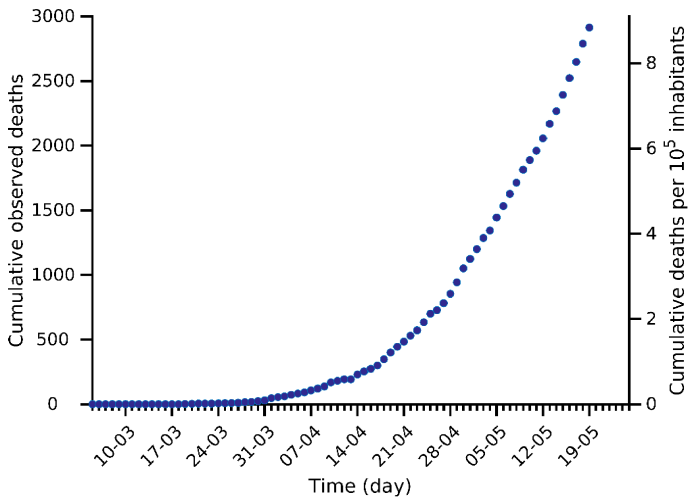


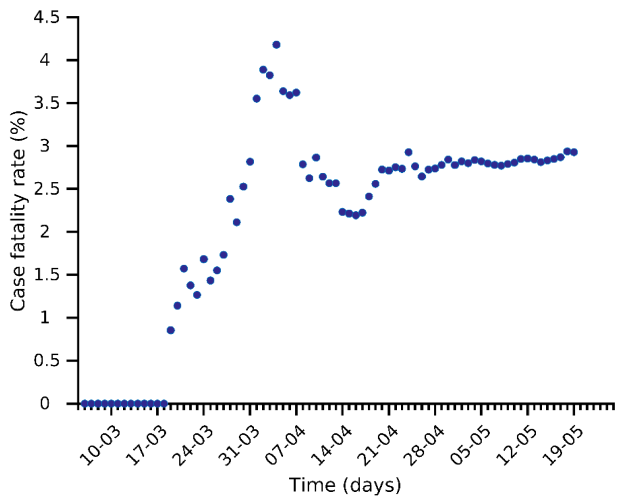




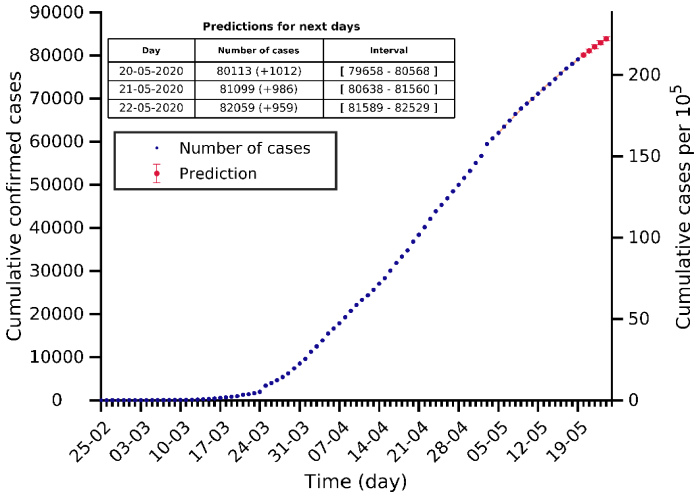


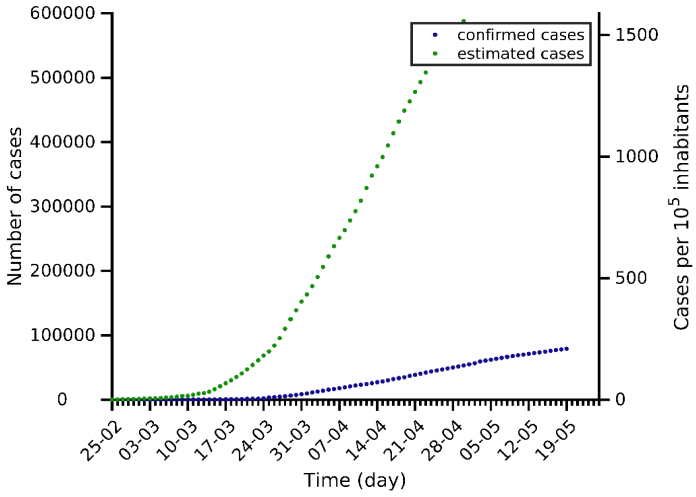


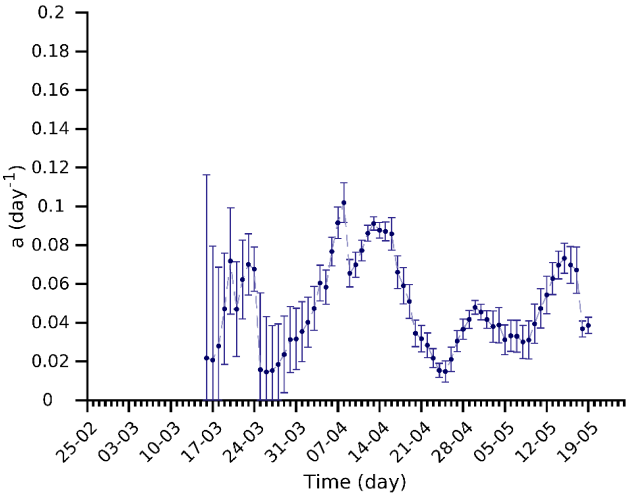


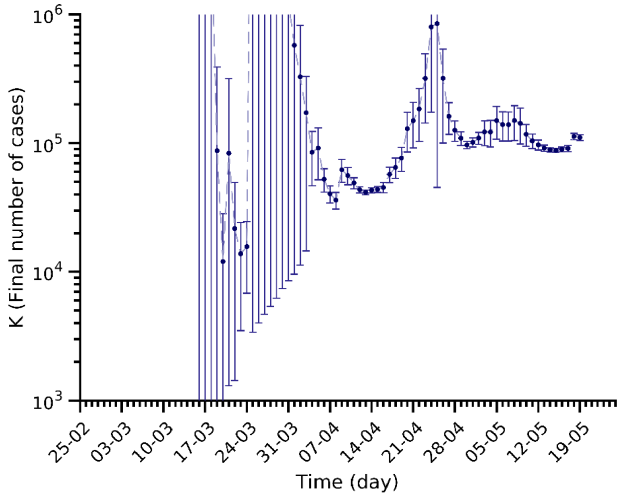


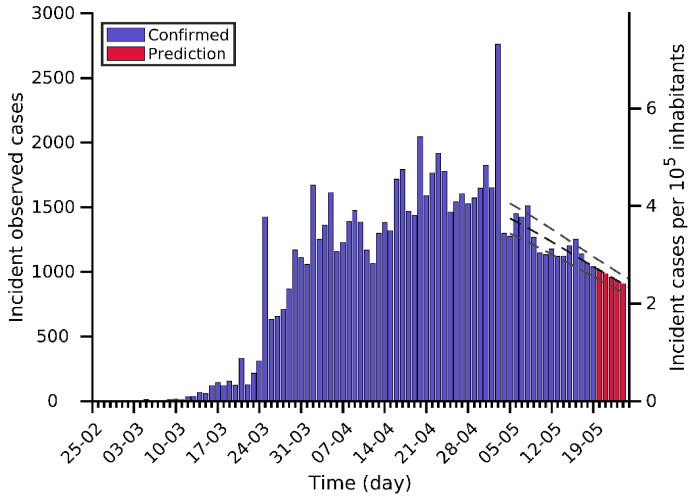
Canada 19-05-2020. Population: 37.7M. Current cumulated incidence: 210/10⁵

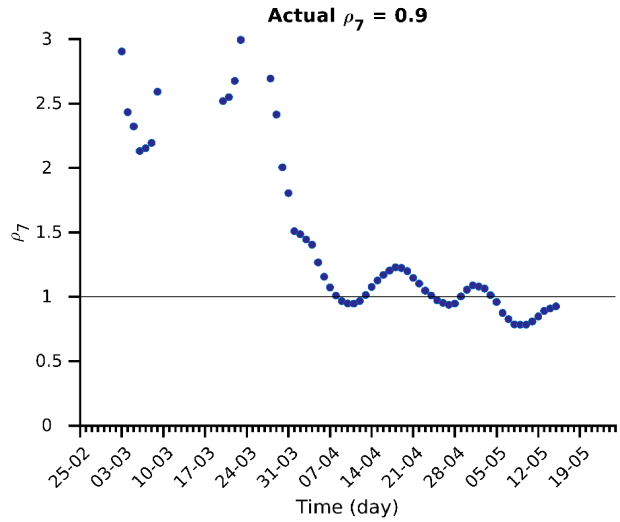


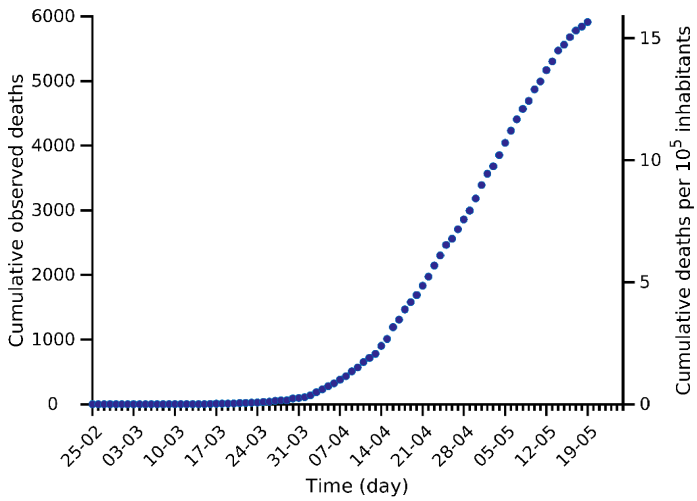


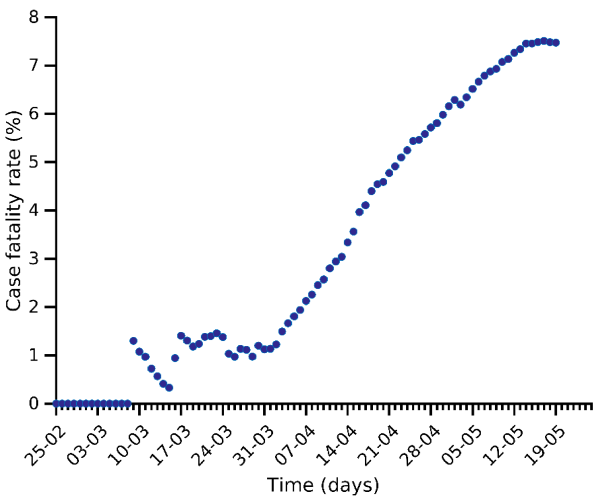




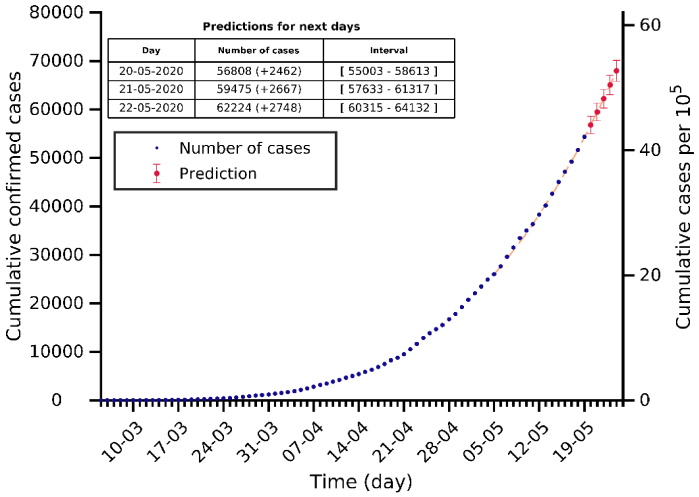


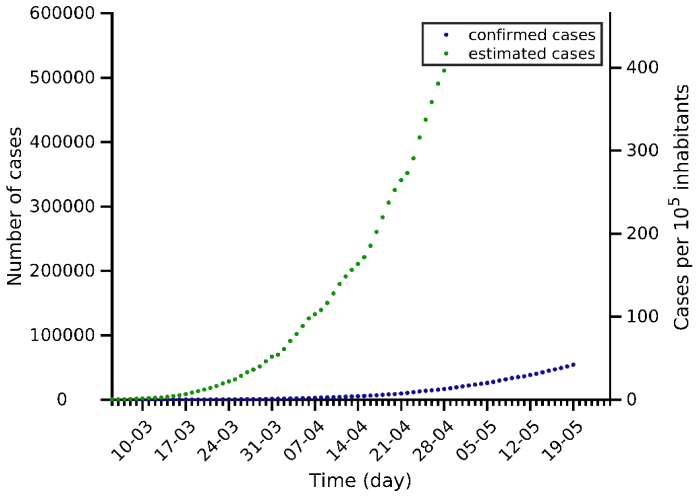


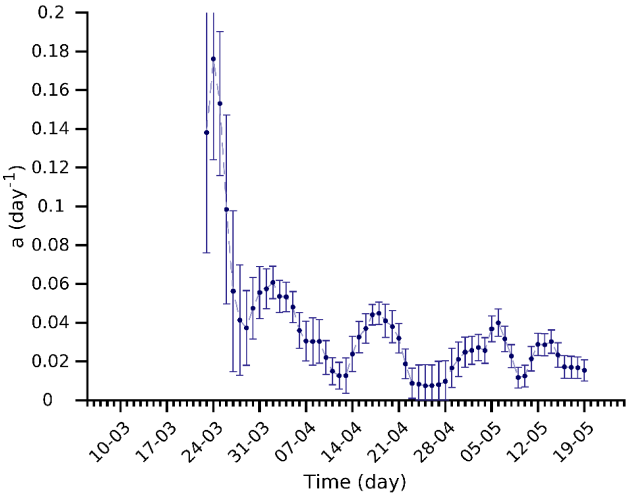


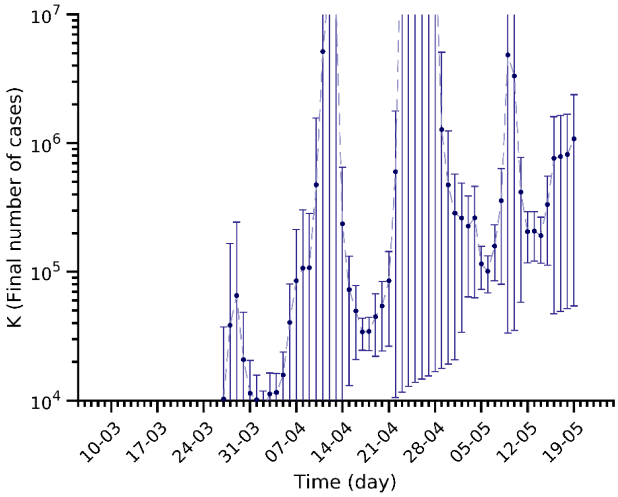


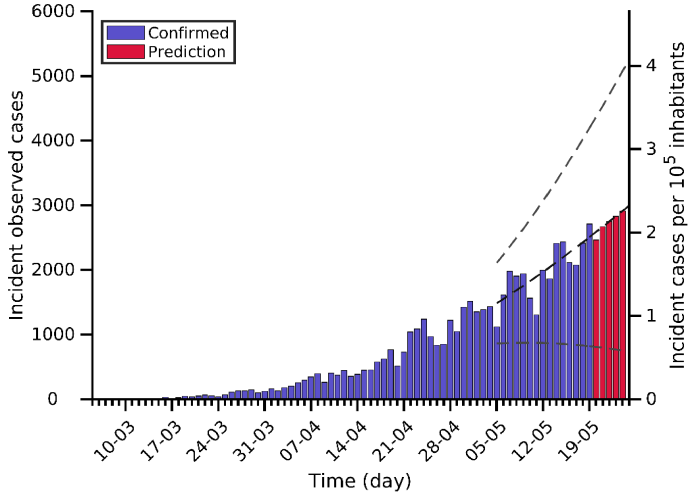
Mexico 19-05-2020. Population: 128.9M. Current cumulated incidence: 42/10⁵

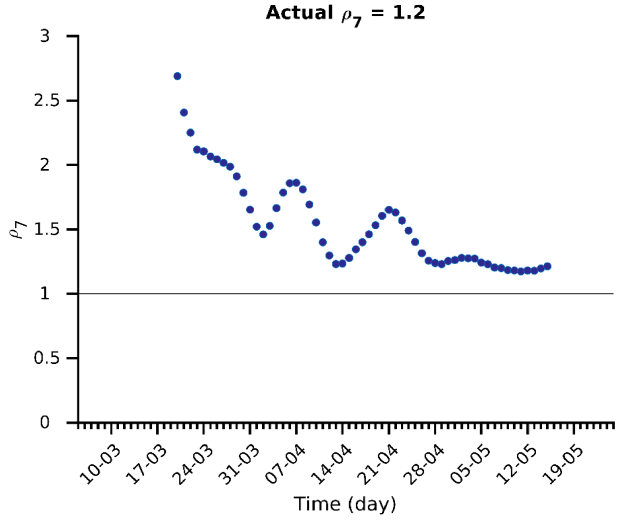


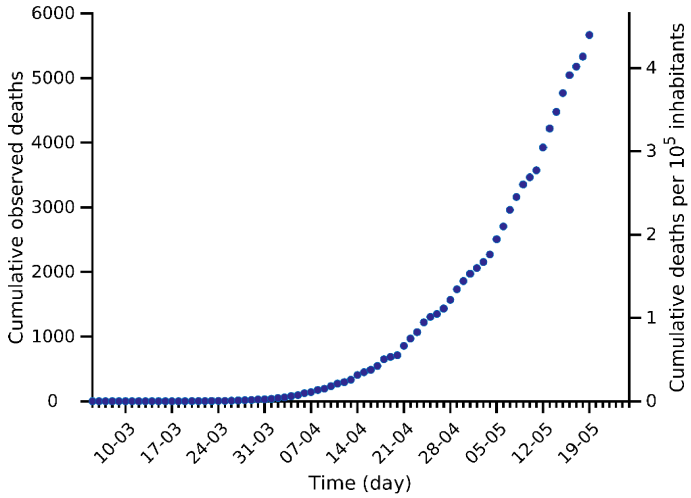


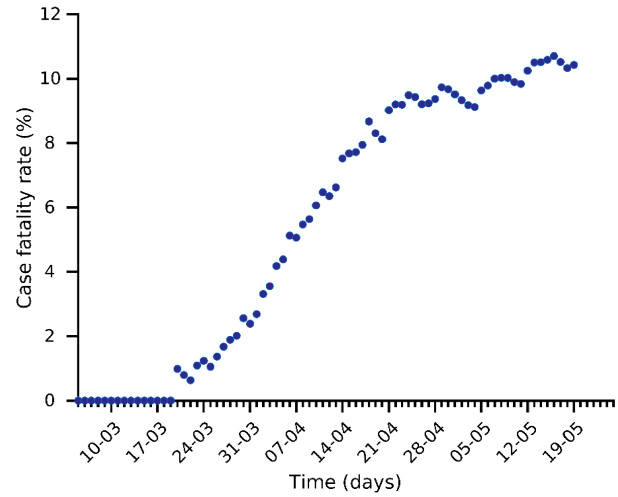




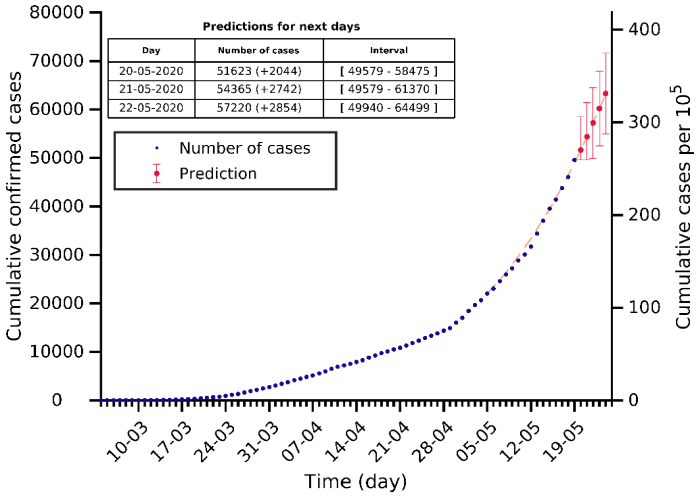


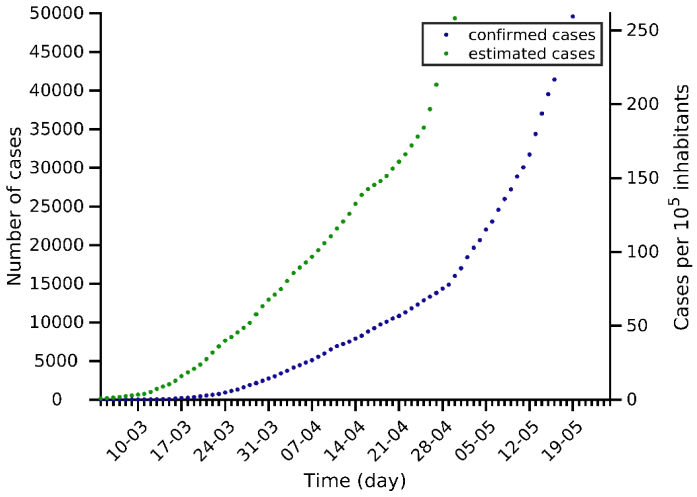


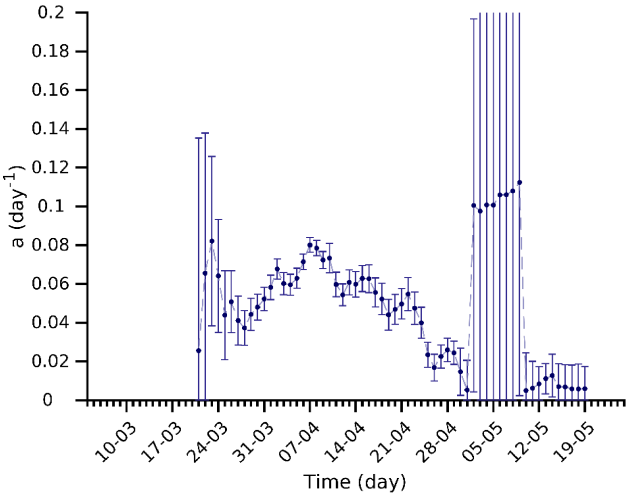


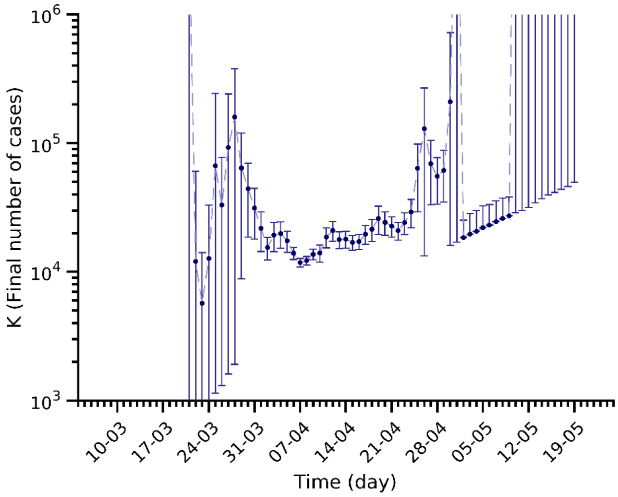


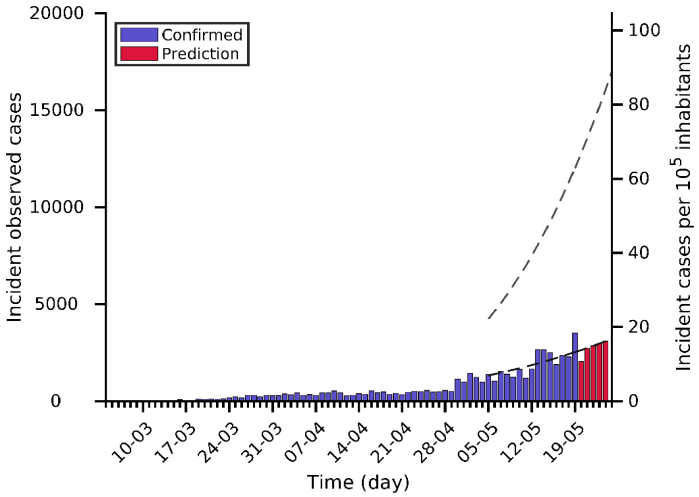
Chile 19-05-2020. Population: 19.1M. Current cumulated incidence: 259/10^5

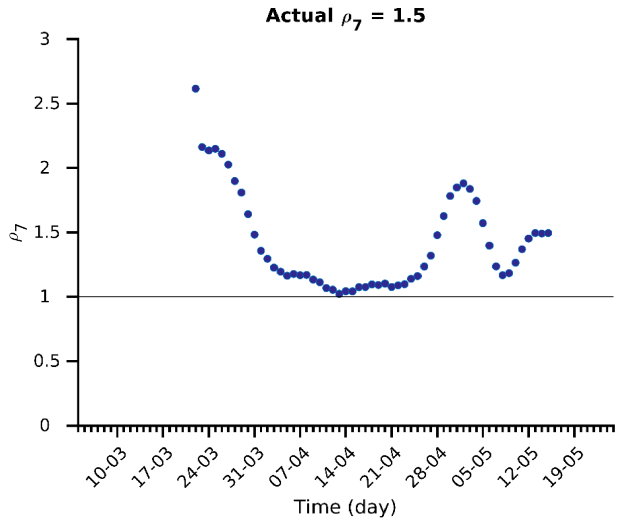


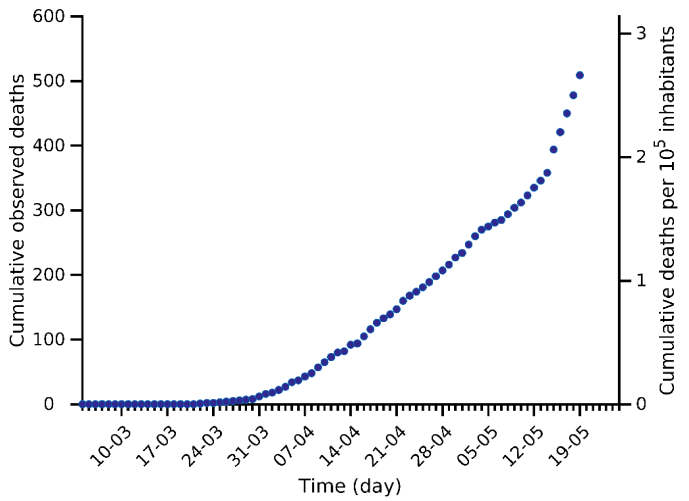


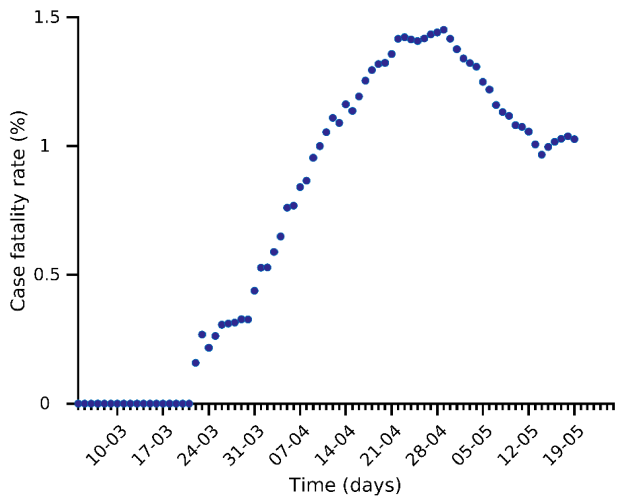




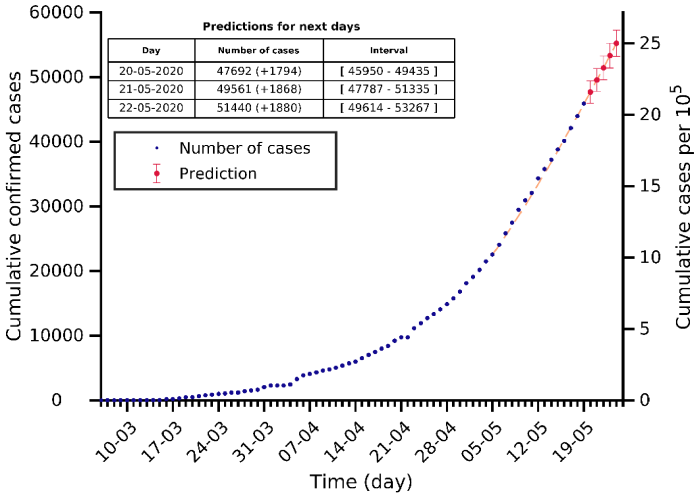


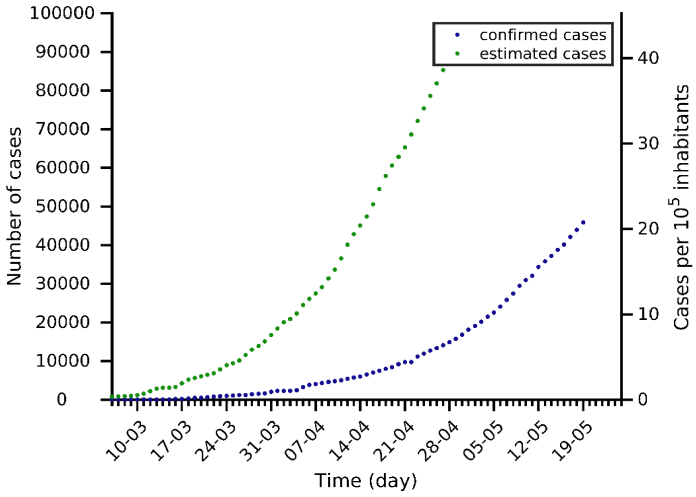


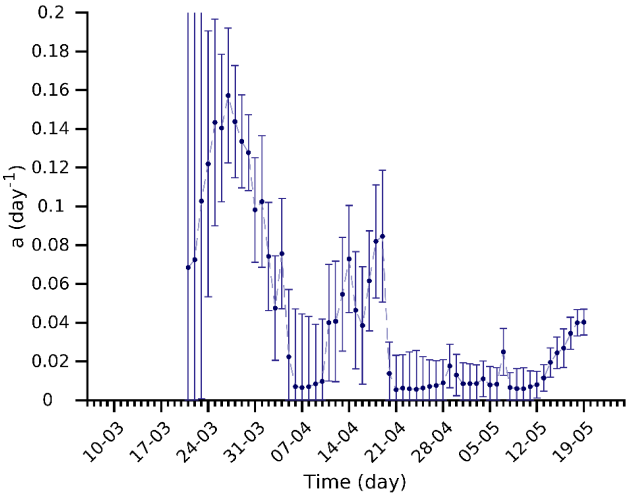


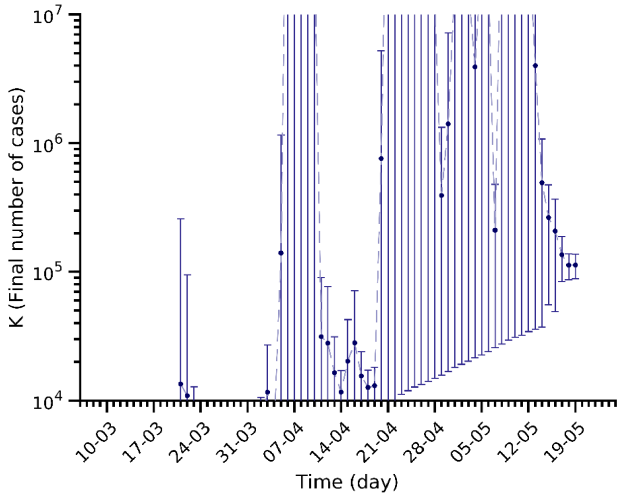


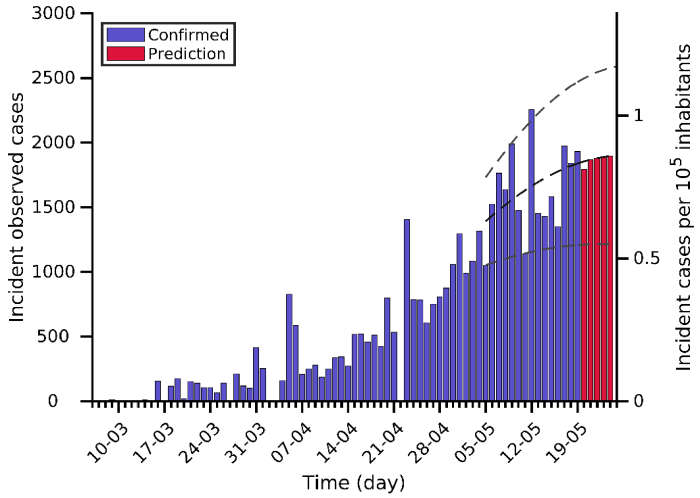
Pakistan 19-05-2020. Population: 220.9M. Current cumulated incidence: $21/10^5$

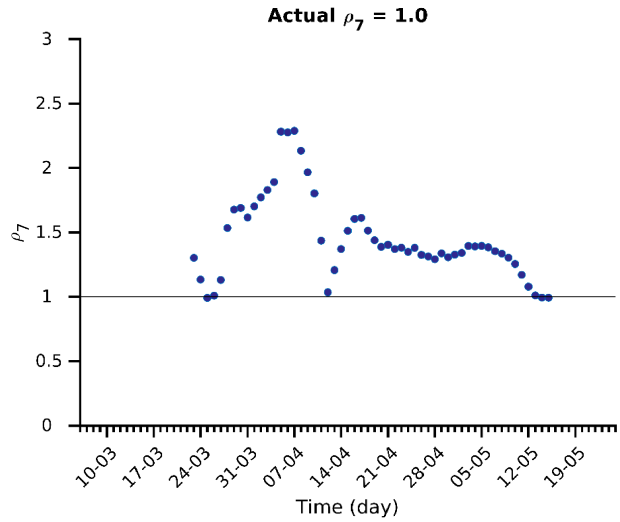


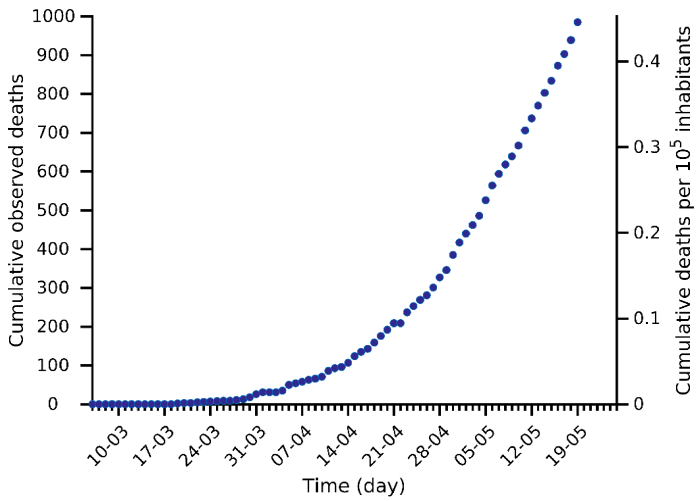


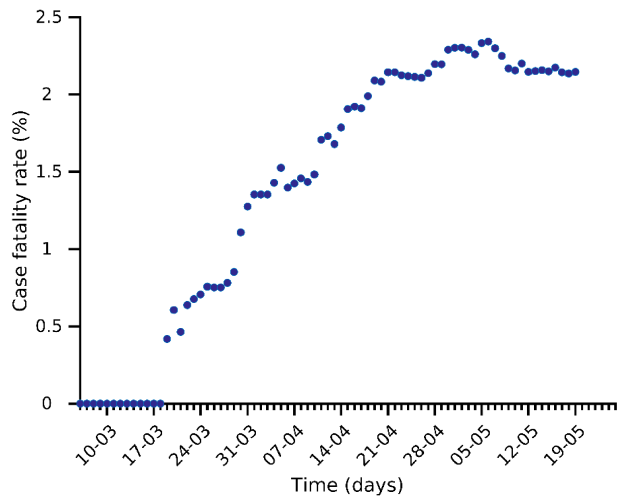




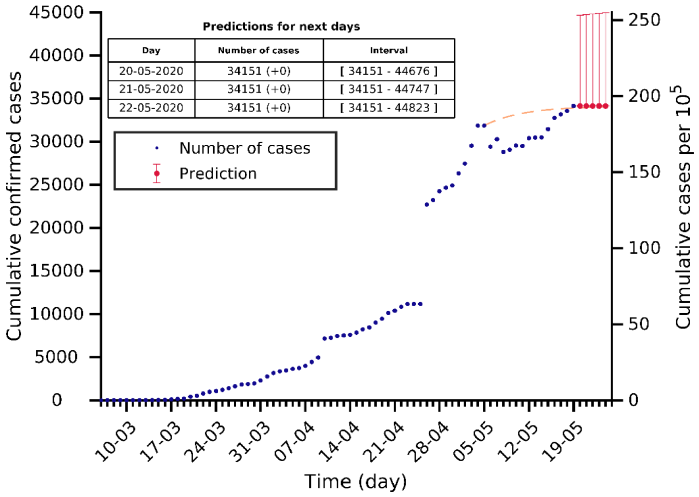


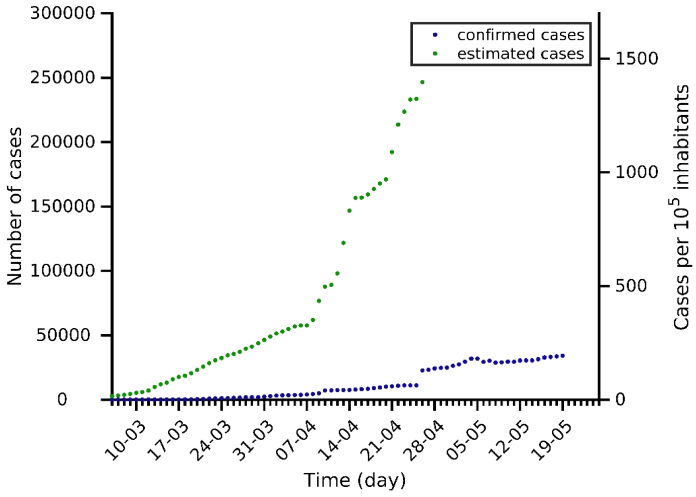


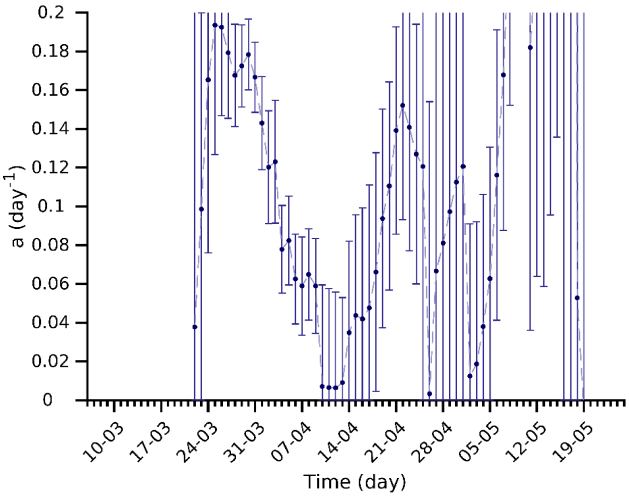


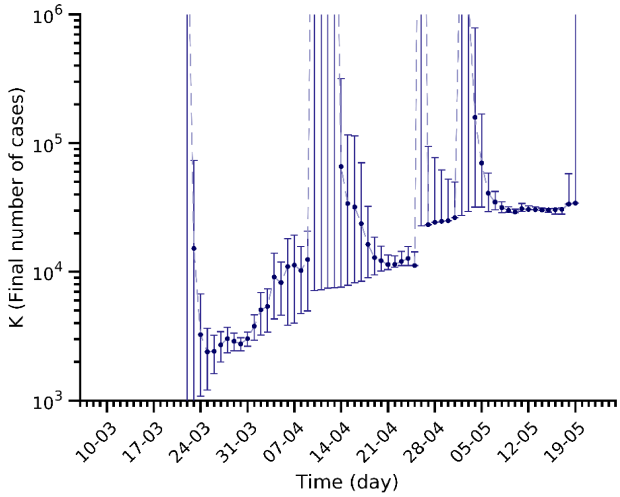


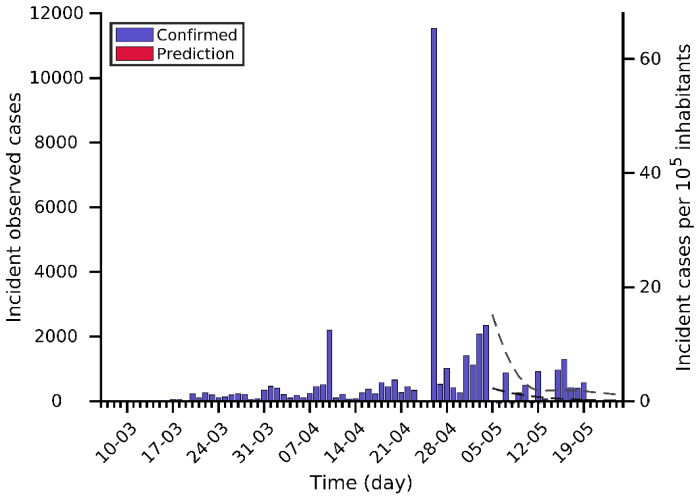
Ecuador 19-05-2020. Population: 17.6M. Current cumulated incidence: 194/10^5

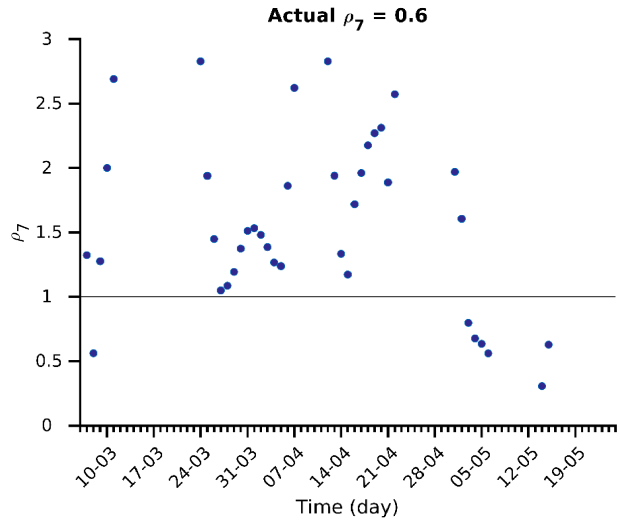


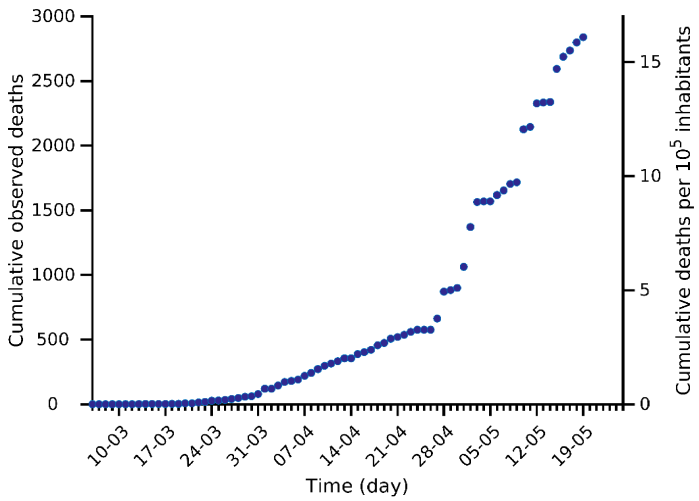


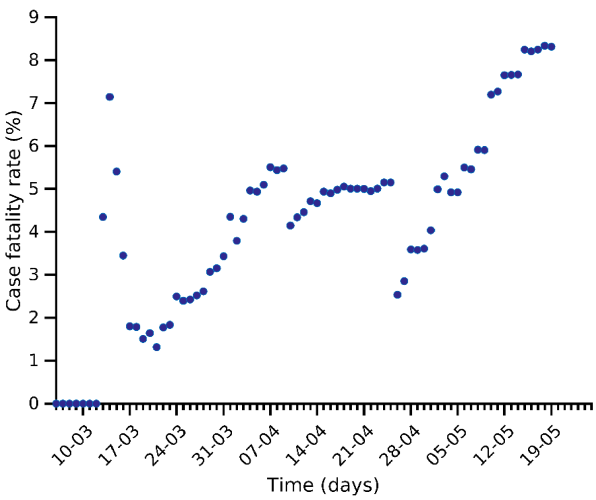




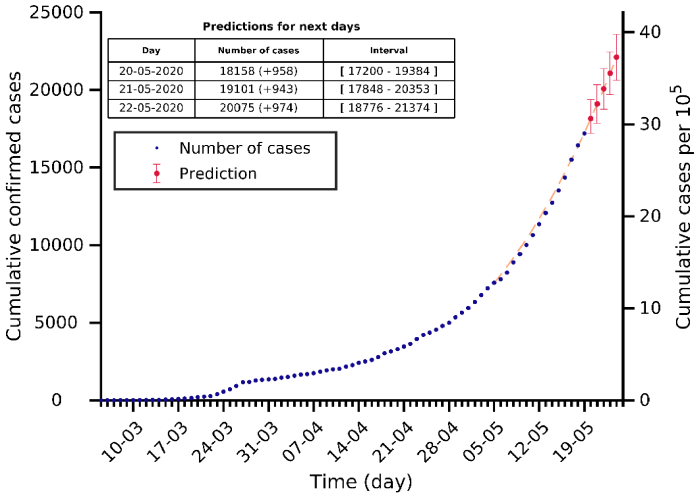


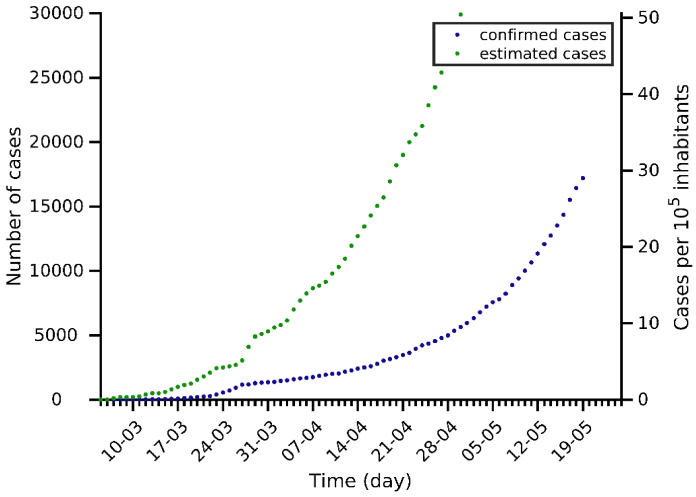


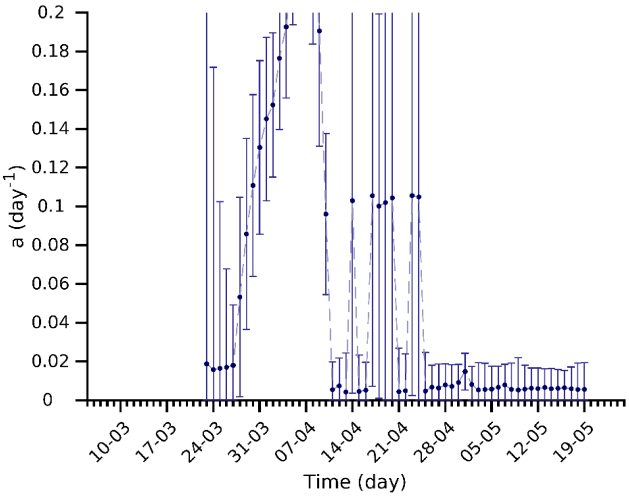


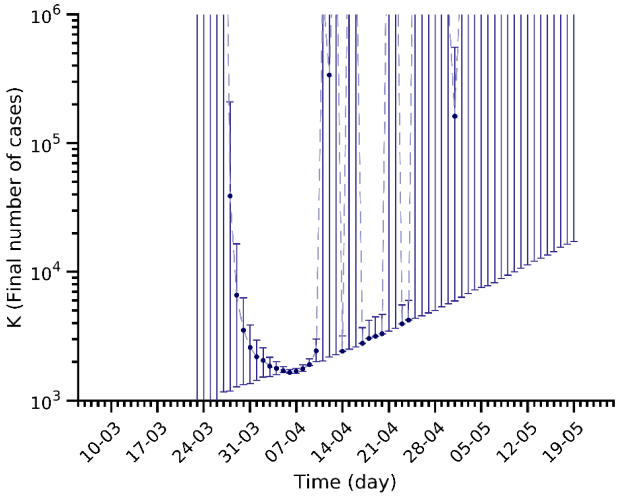


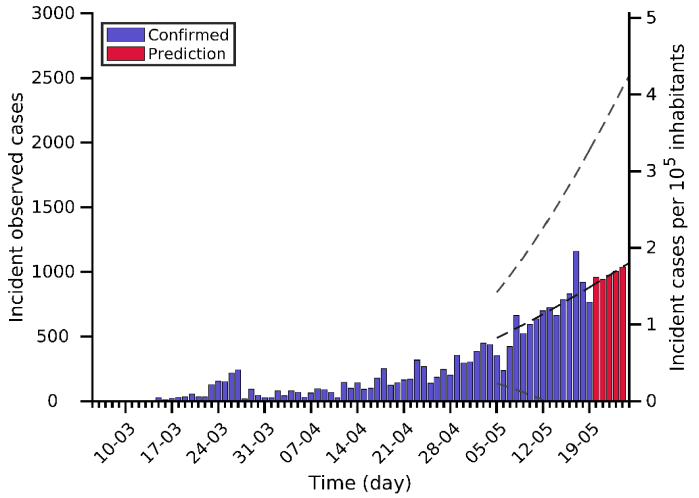
South Africa 19-05-2020. Population: 59.3M. Current cumulated incidence: 29/10^5

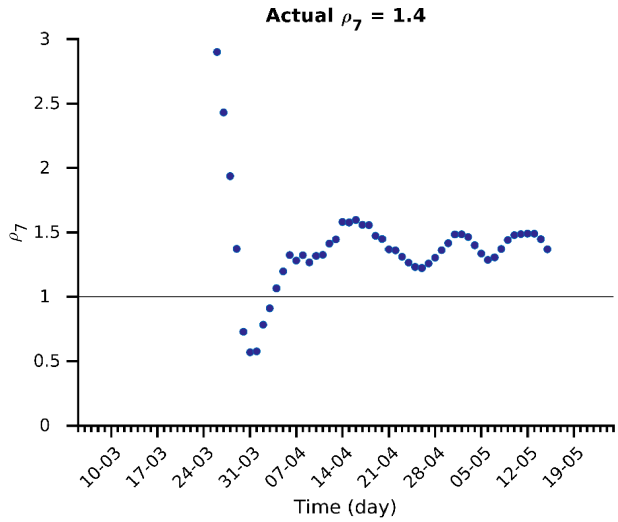


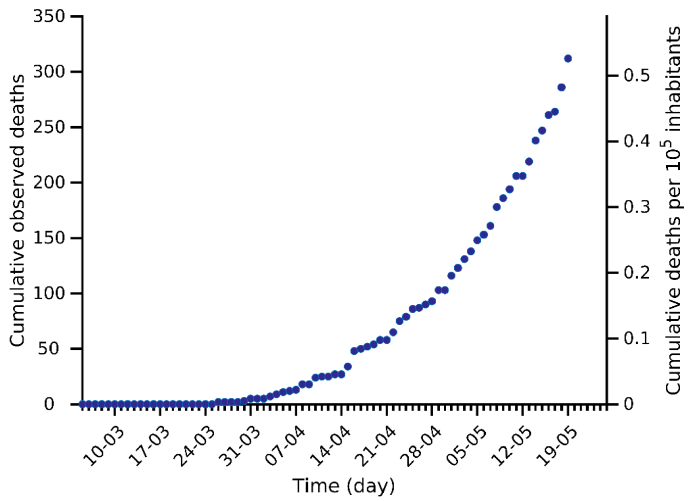


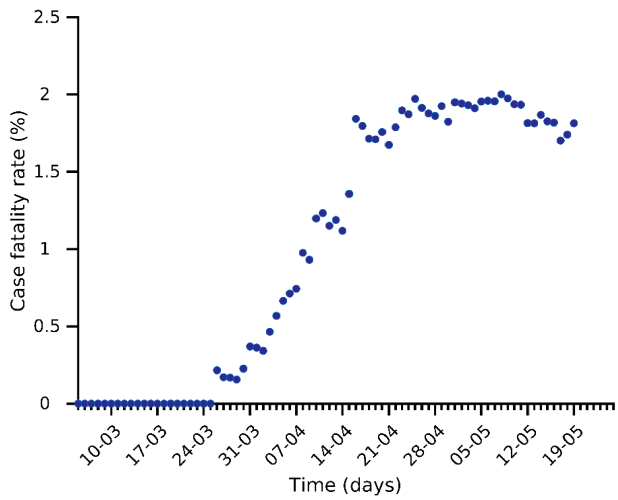




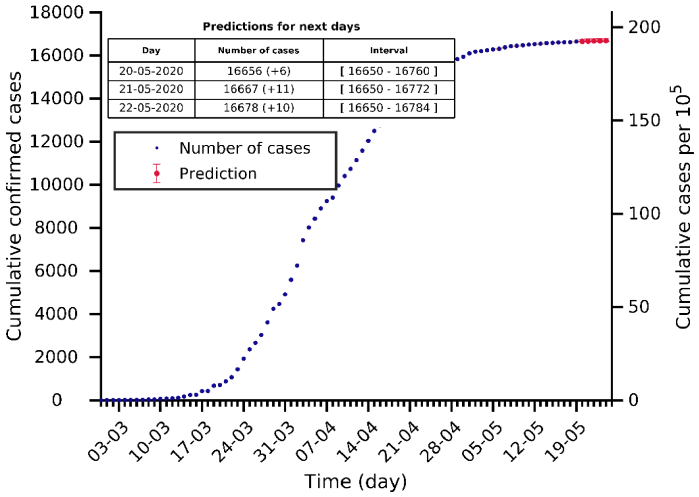


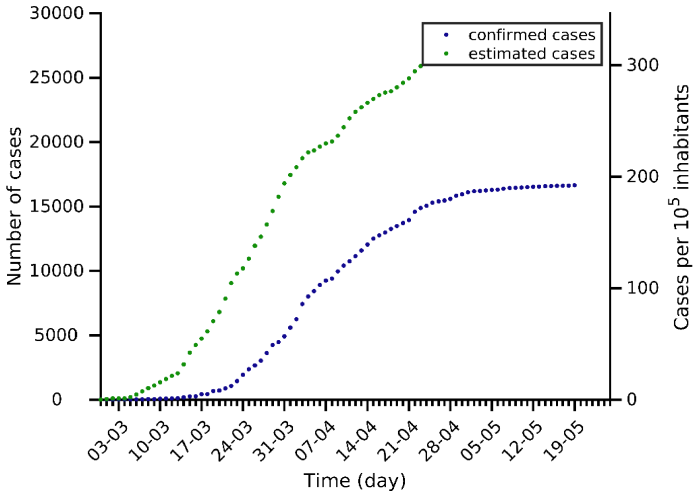


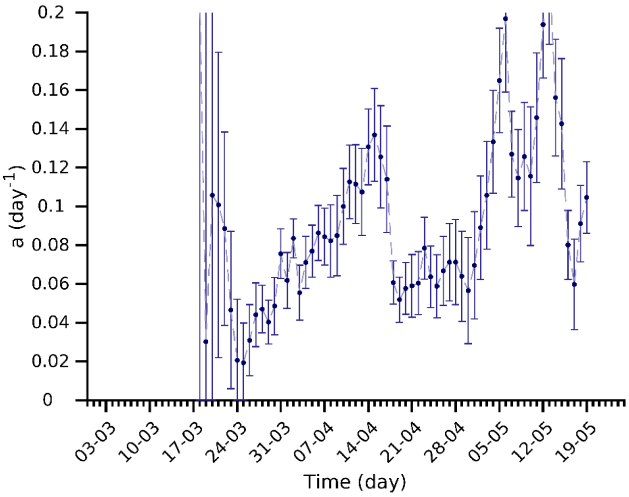


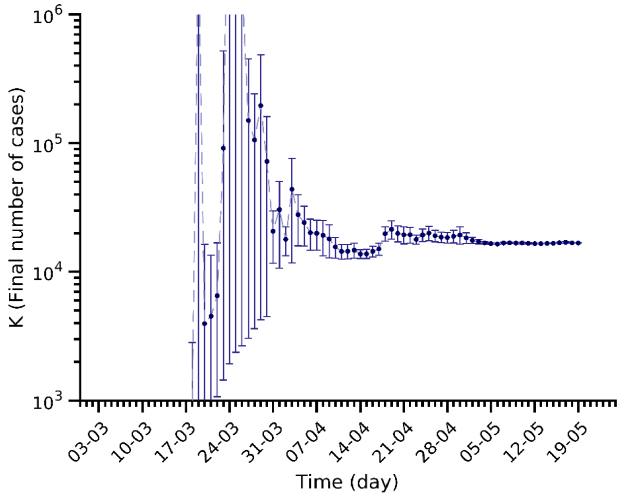


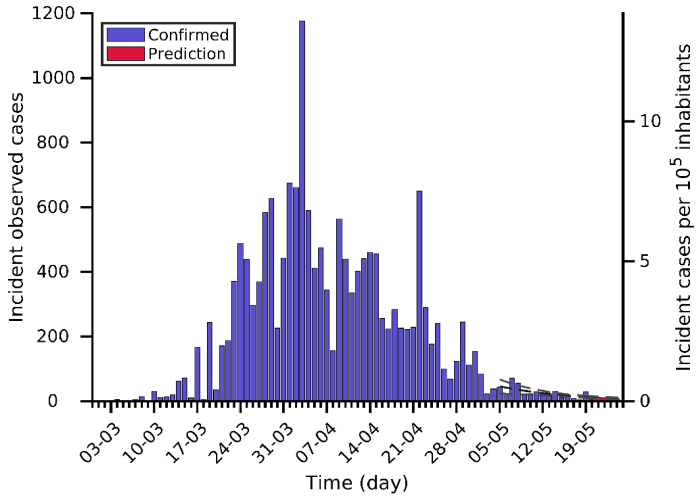
Israel 19-05-2020. Population: 8.7M. Current cumulated incidence: 192/10^5

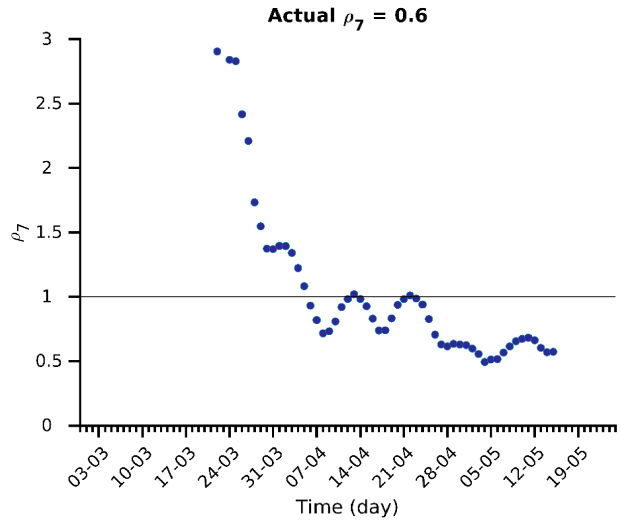


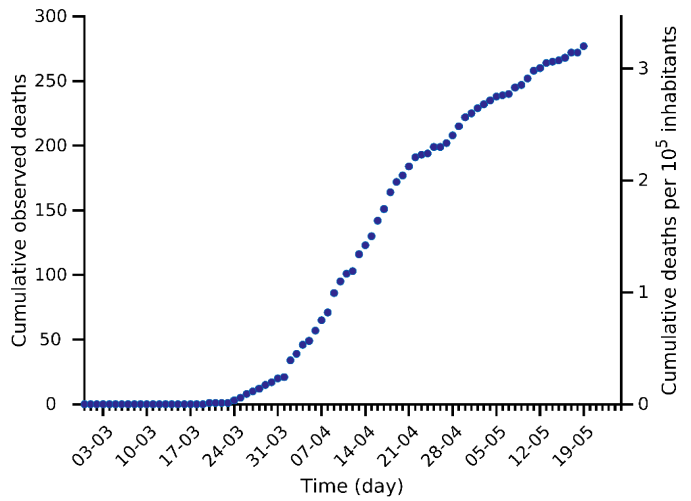


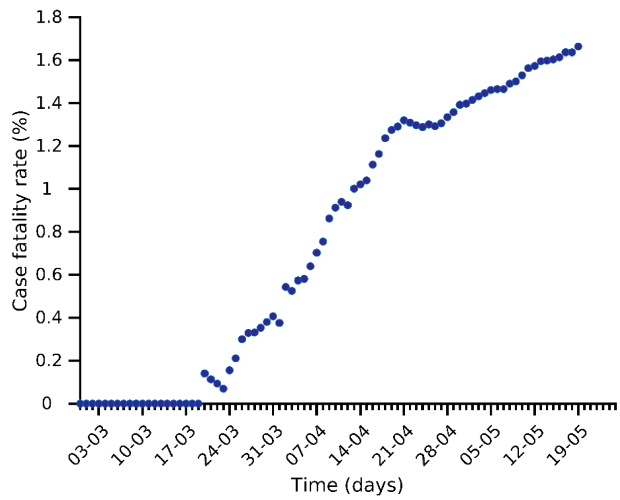




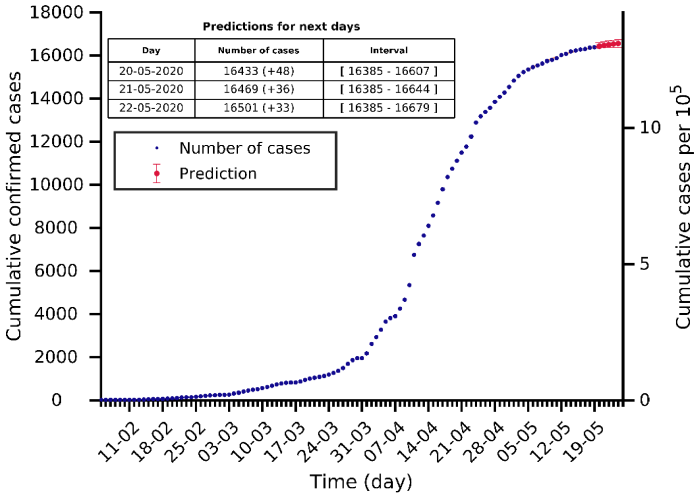


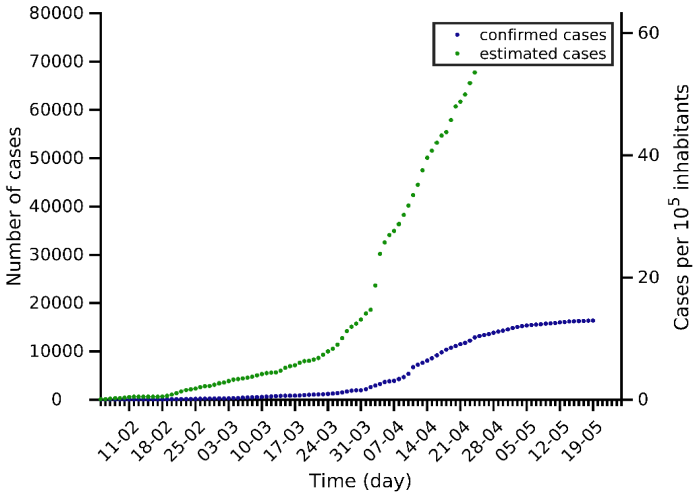


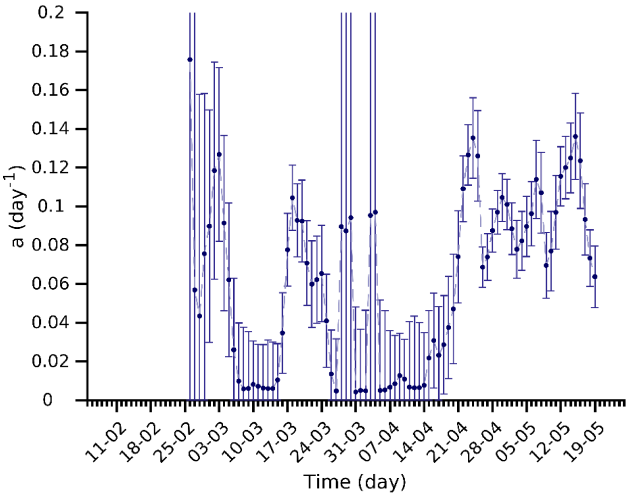


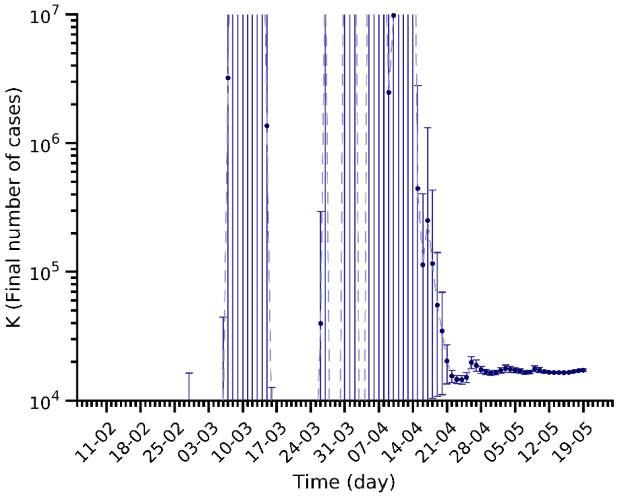


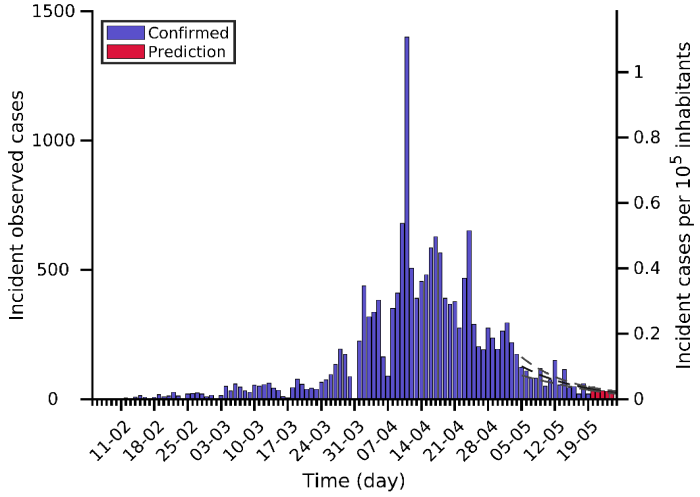
Japan 19-05-2020. Population: 126.5M. Current cumulated incidence: $13/10^5$

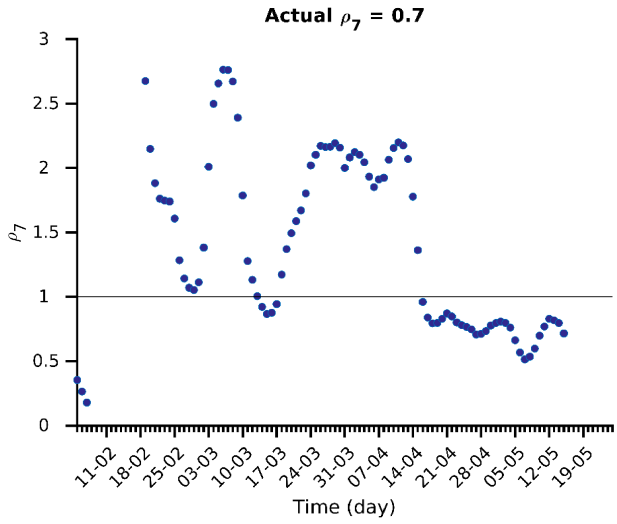


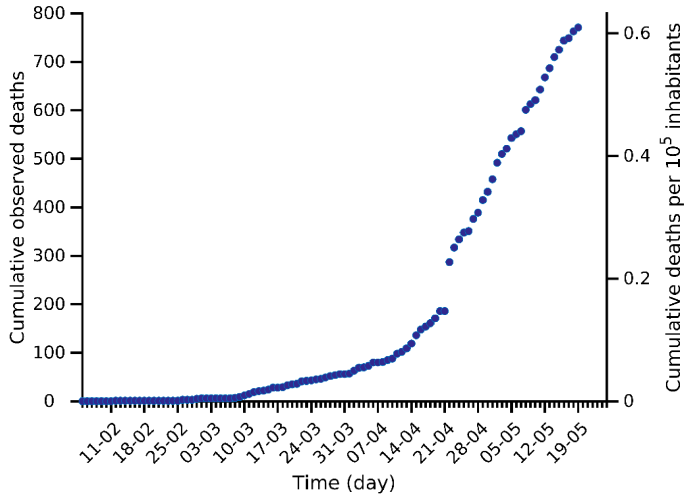


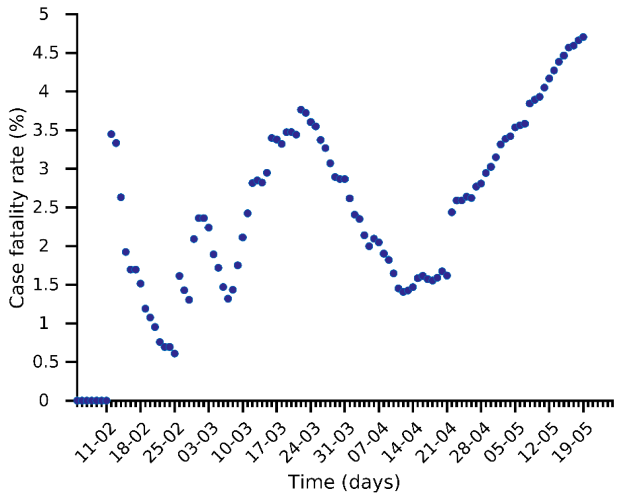




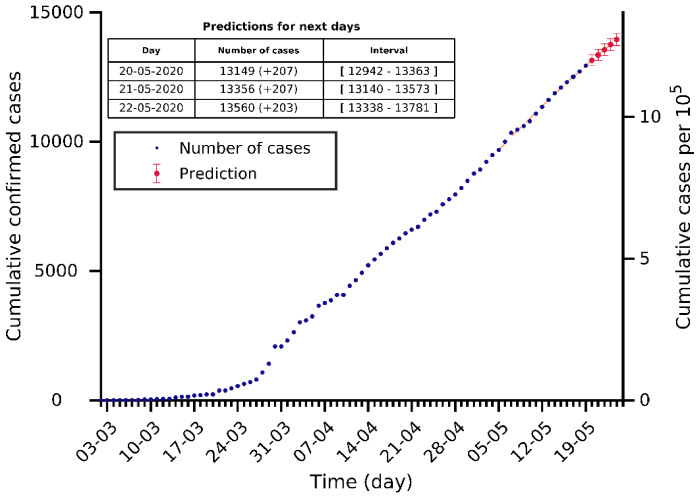


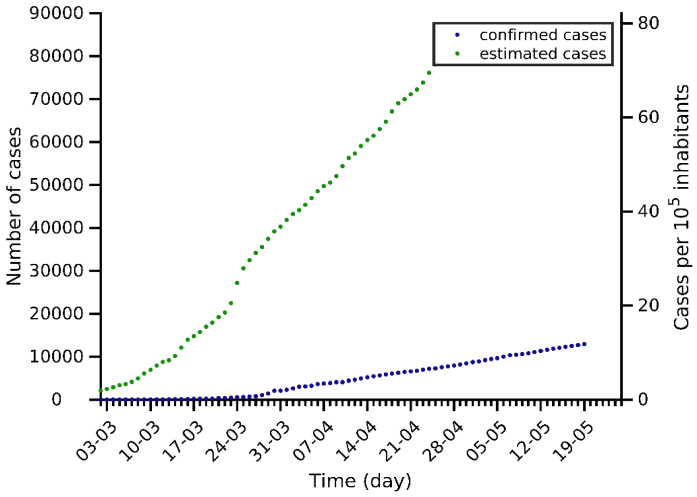


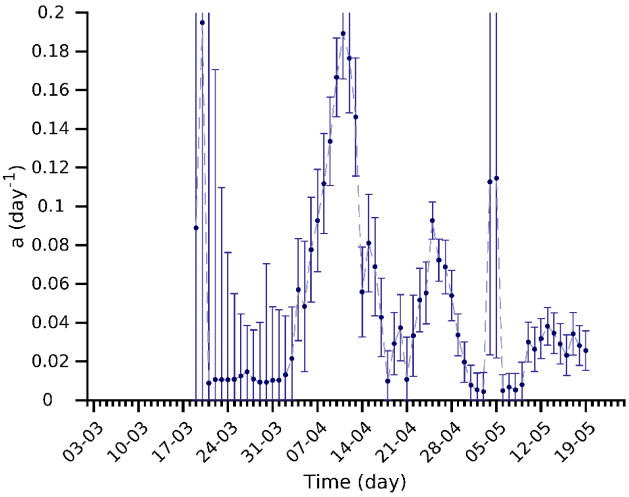


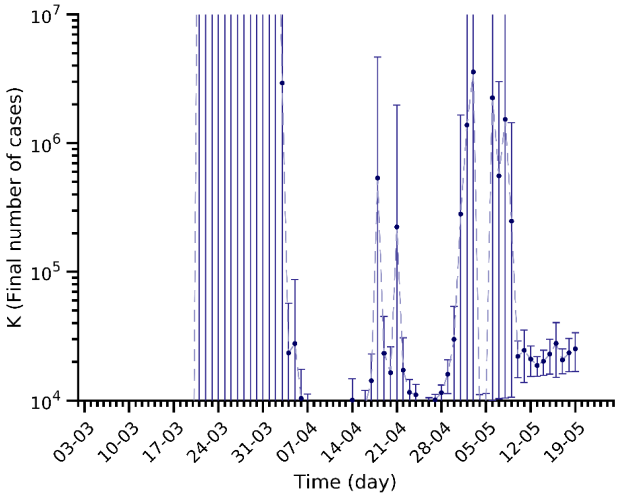


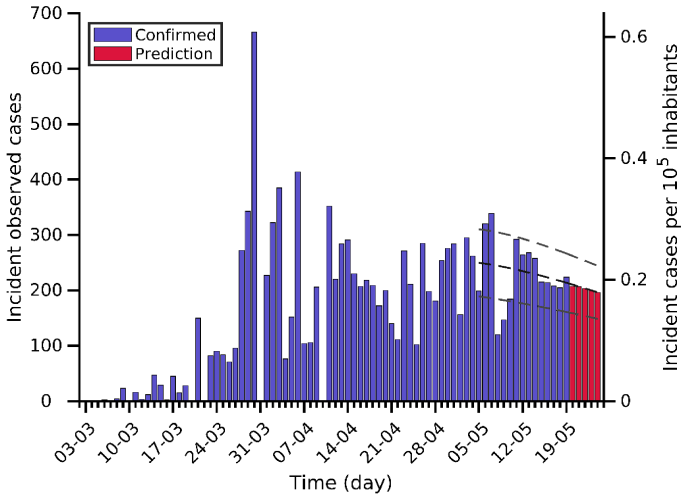
Philippines 19-05-2020. Population: 109.6M. Current cumulated incidence: $12/10^5$

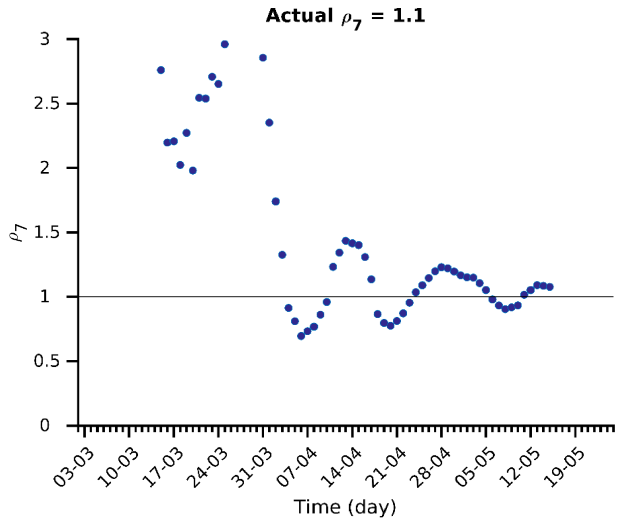


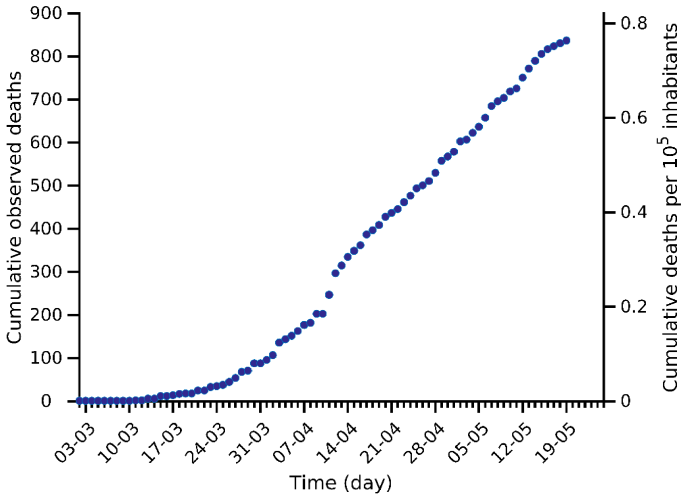


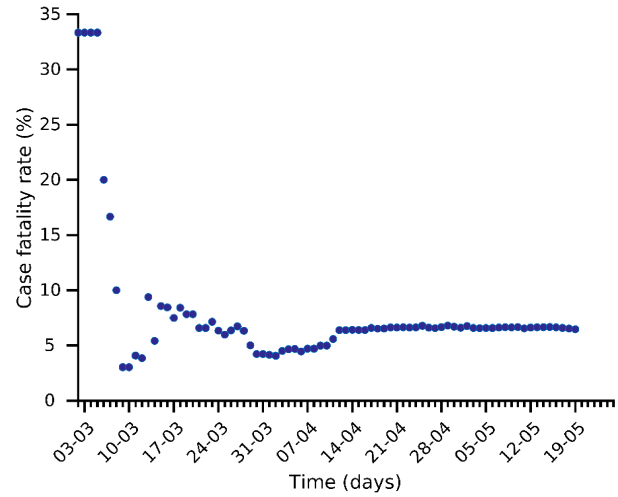




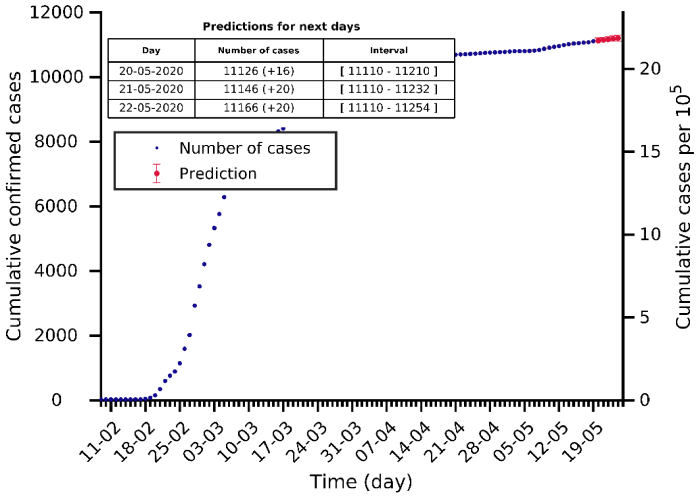


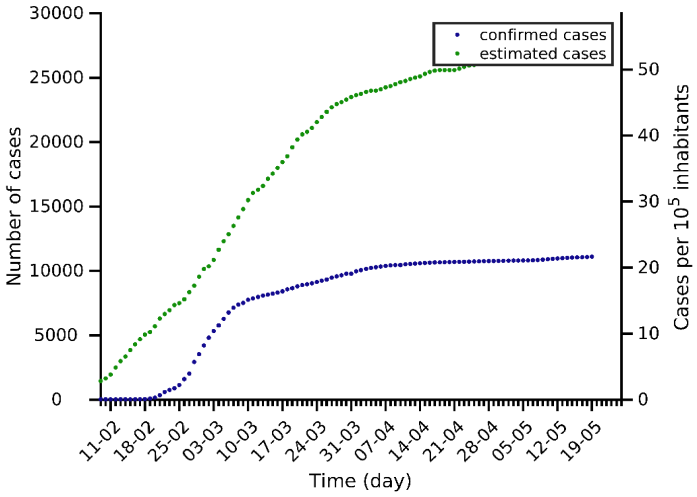


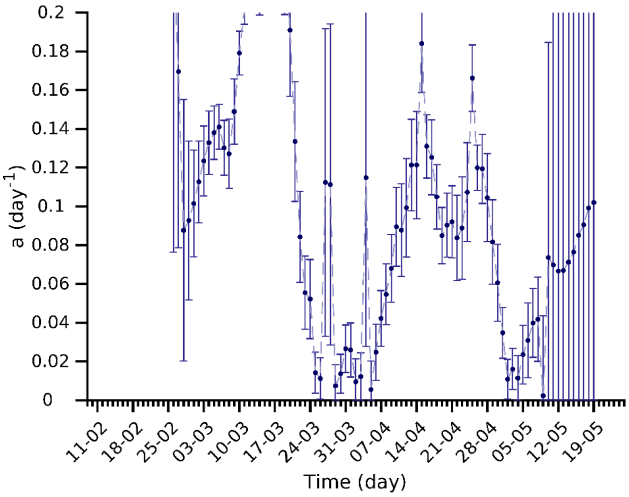


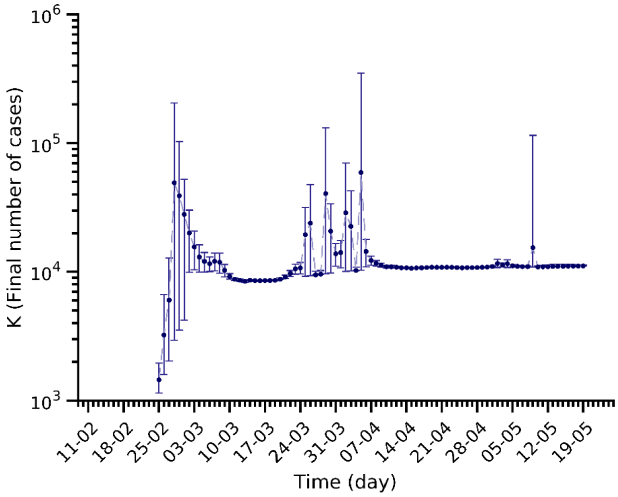


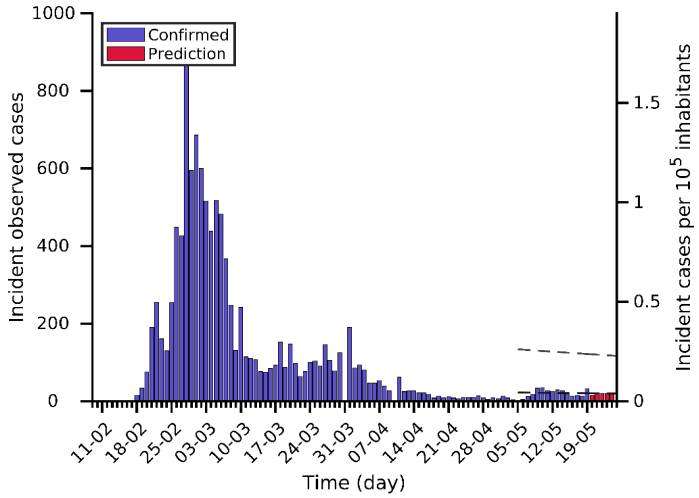
South Korea 19-05-2020. Population: 51.3M. Current cumulated incidence: $22/10^5$

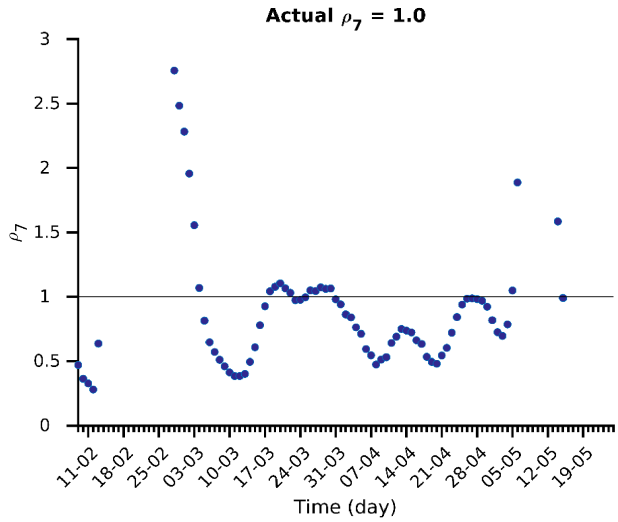


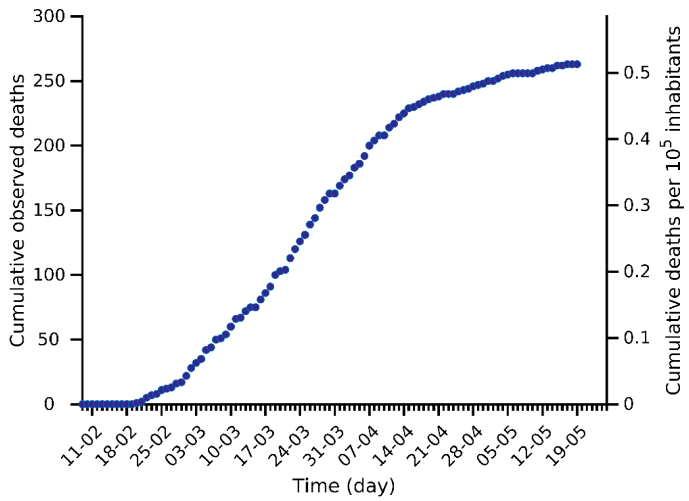


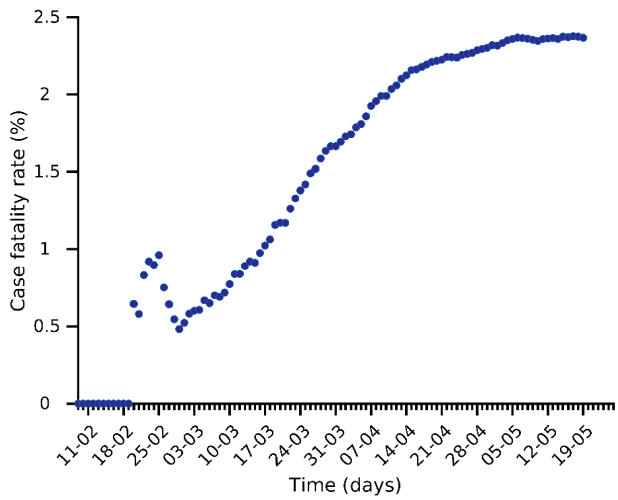




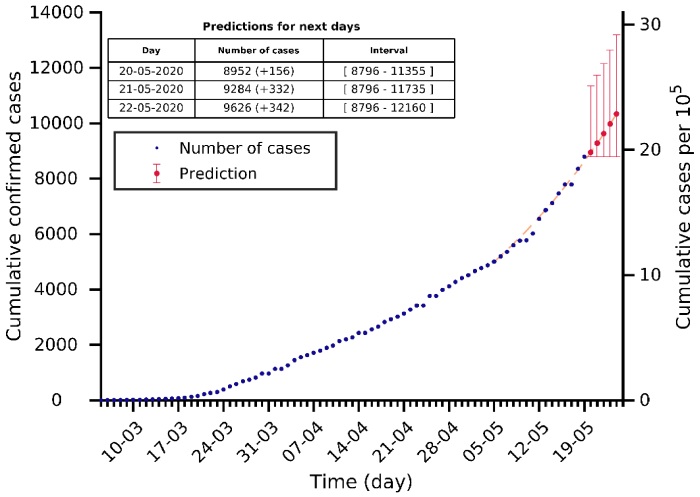


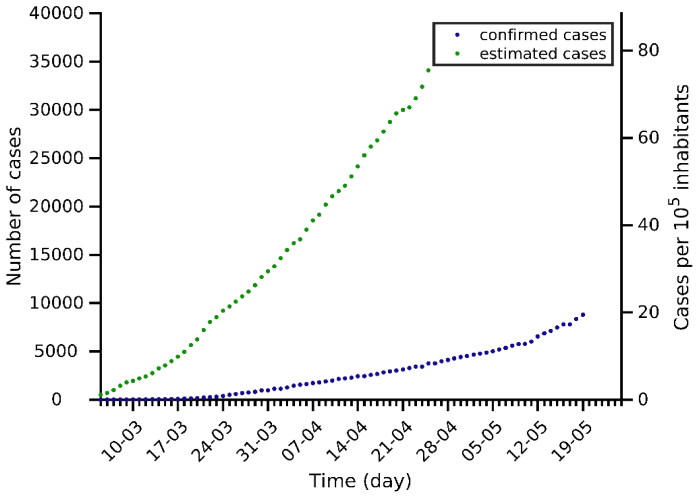


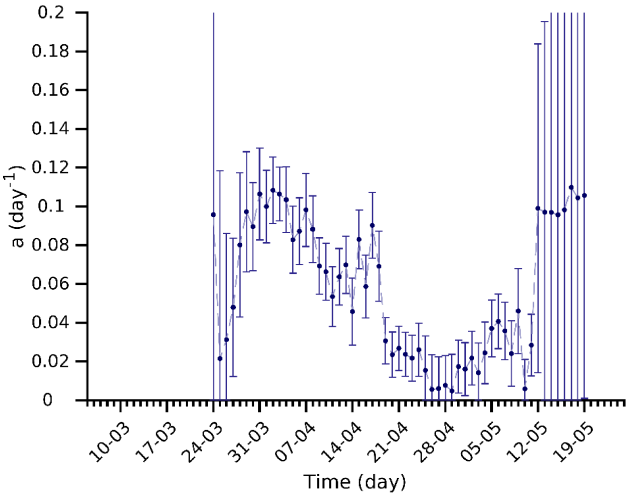


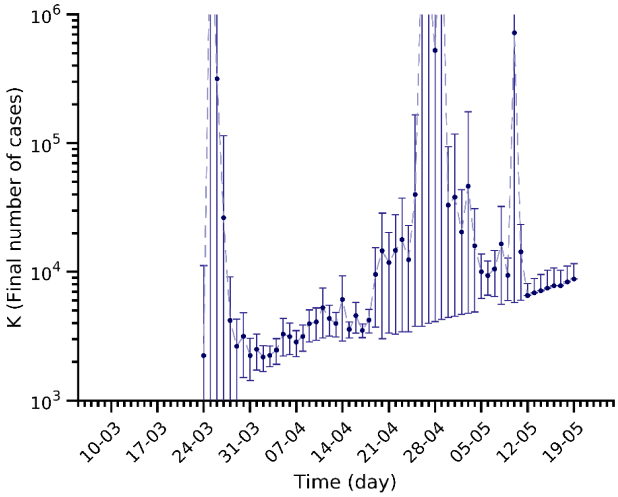


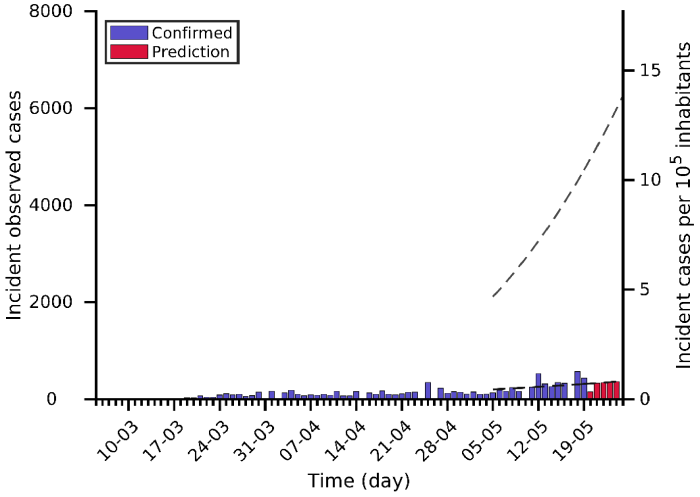
Argentina 19-05-2020. Population: 45.2M. Current cumulated incidence: 19/10^5

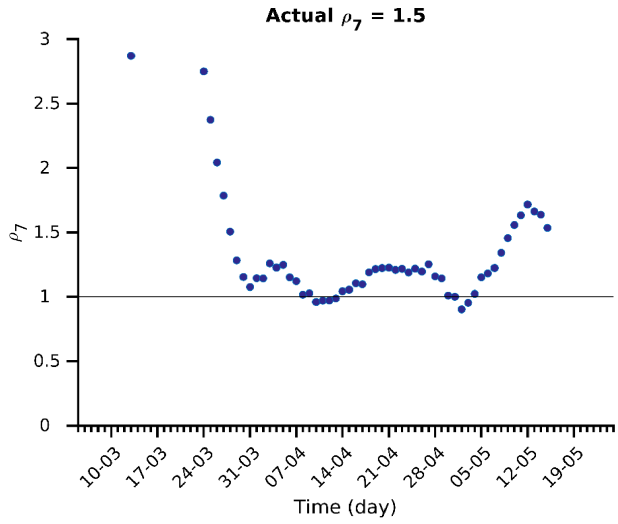


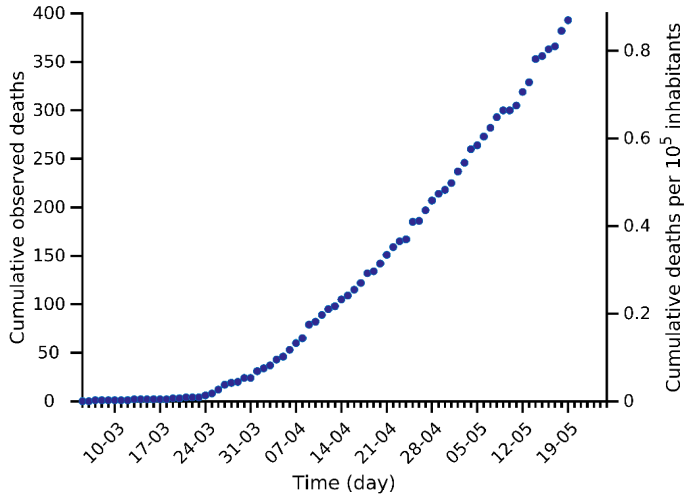


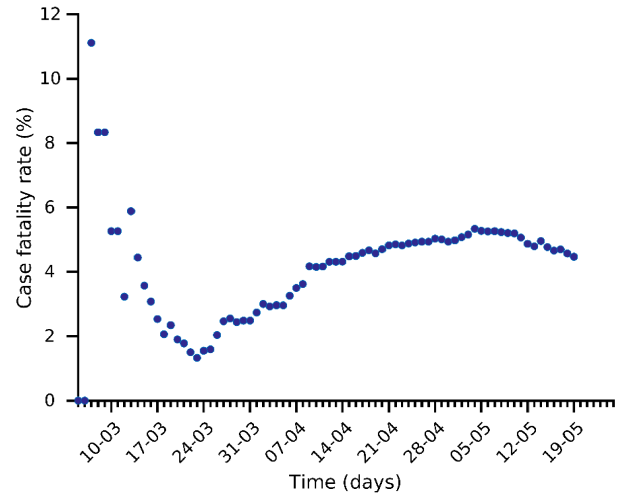




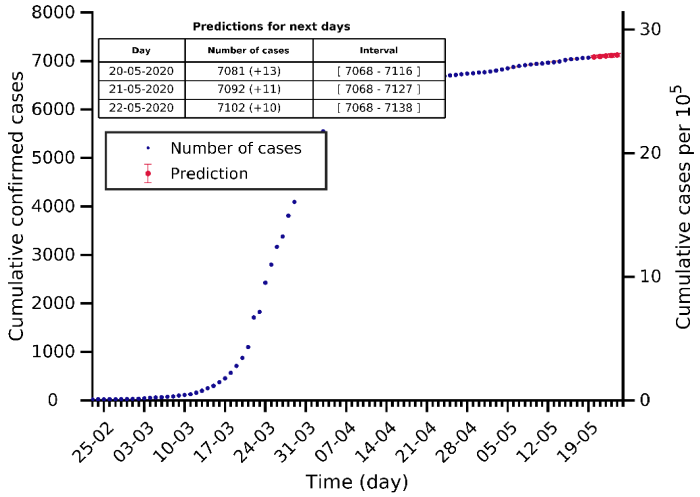


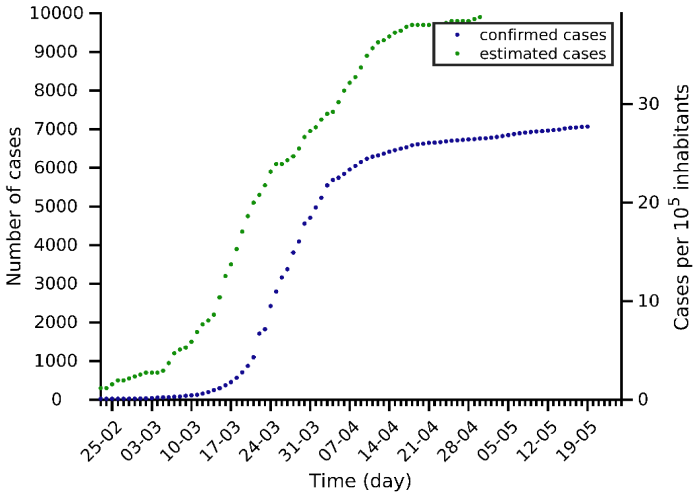


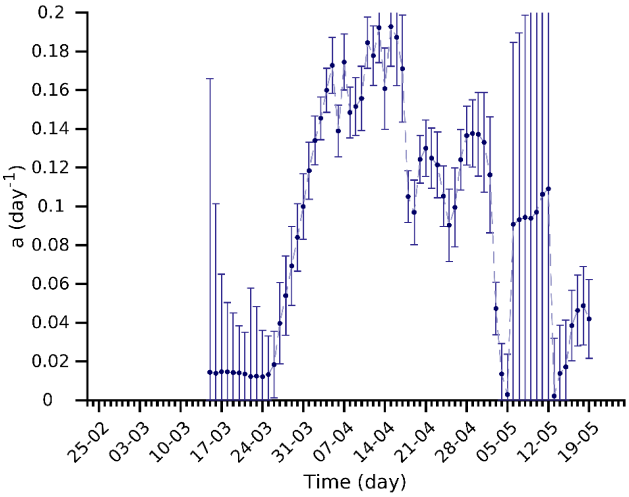


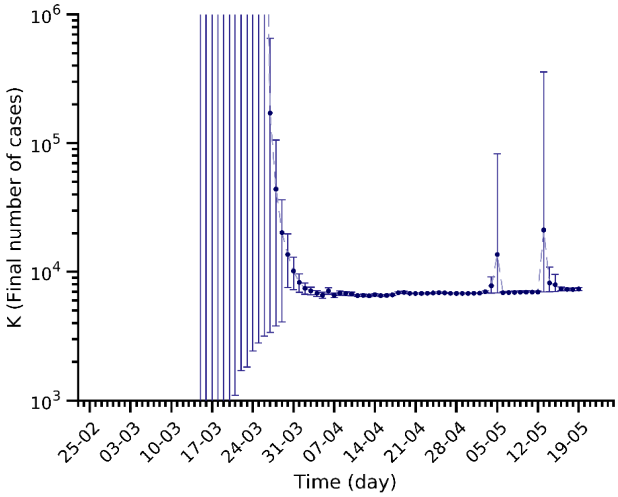


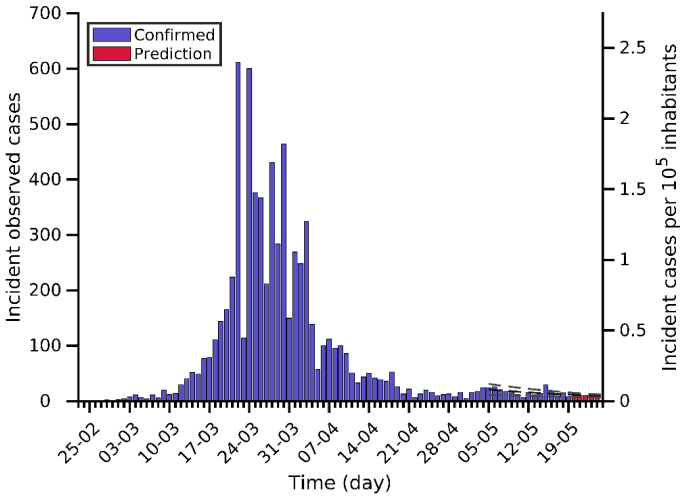
Australia 19-05-2020. Population: 25.5M. Current cumulated incidence: $28/10^5$

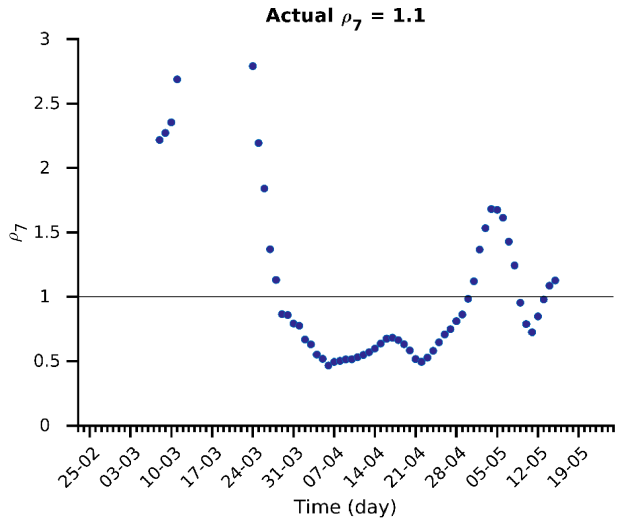


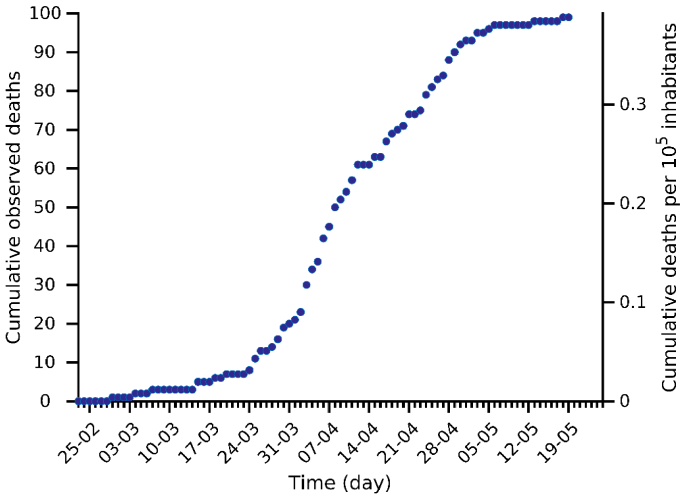


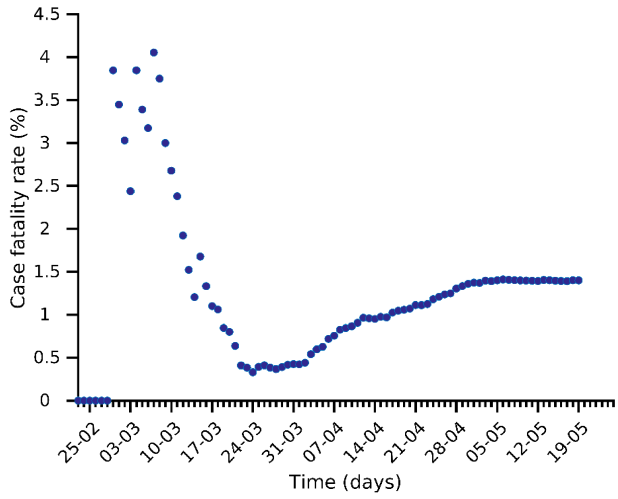




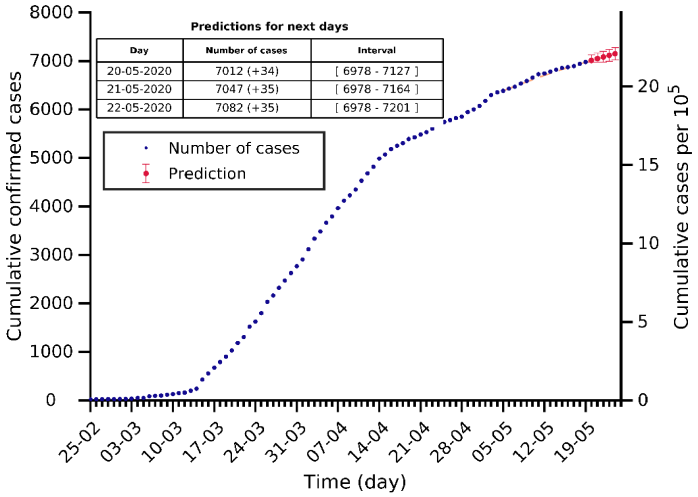


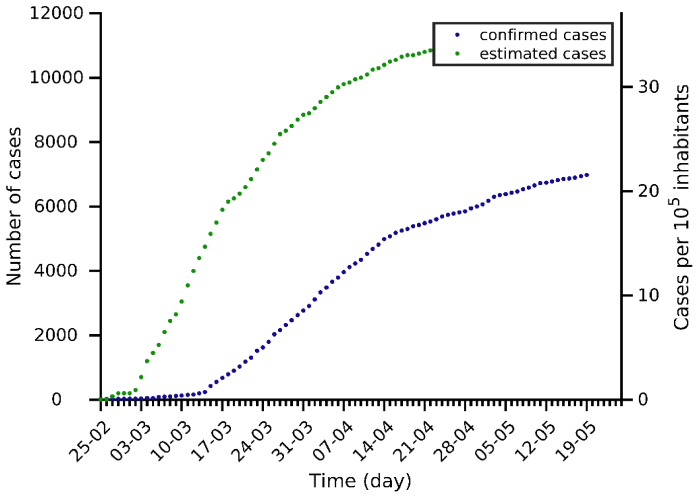


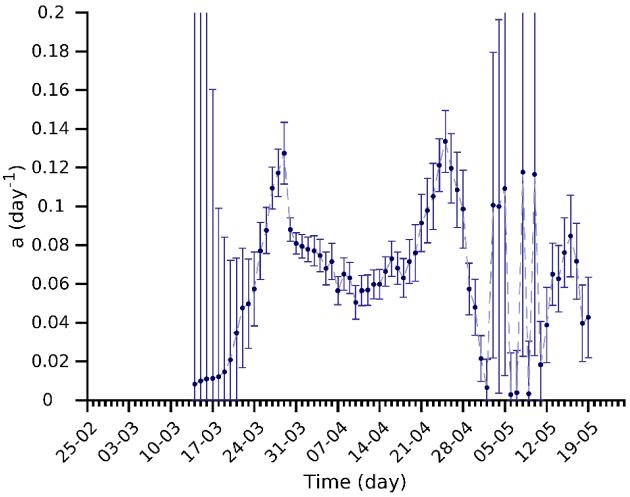


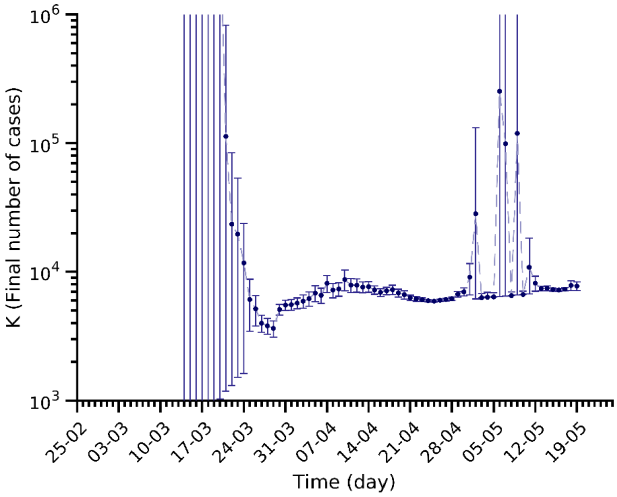


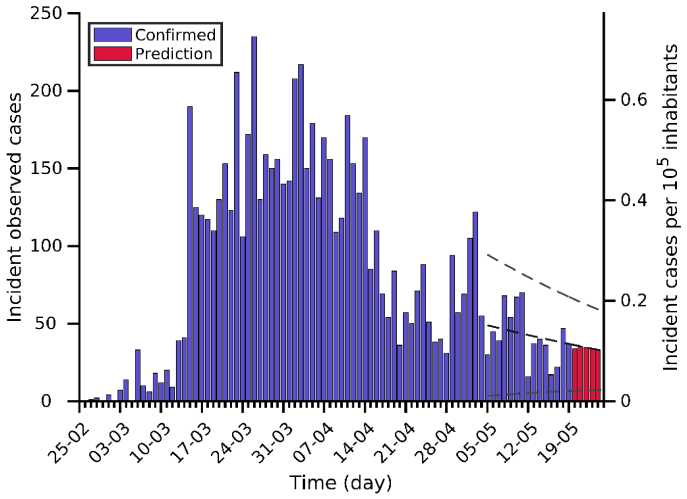
Malaysia 19-05-2020. Population: 32.4M. Current cumulated incidence: $22/10^5$

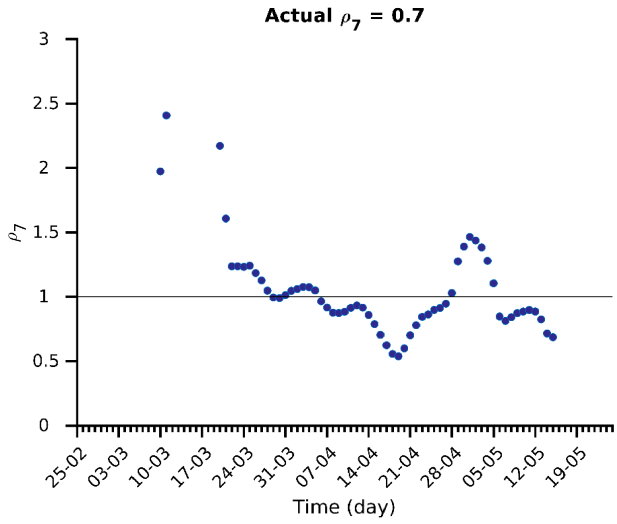


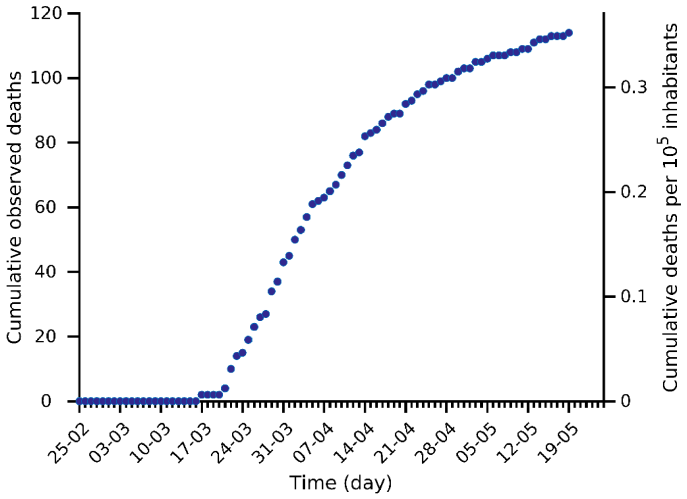


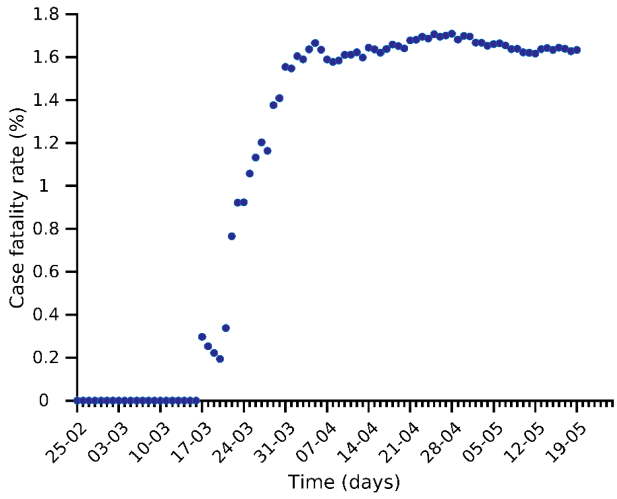




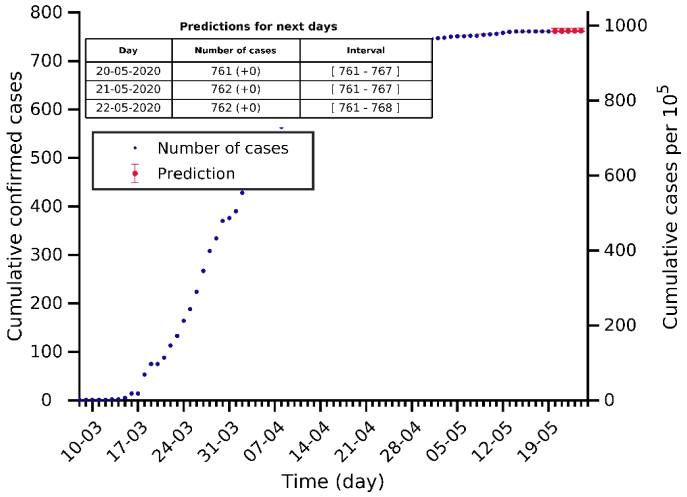


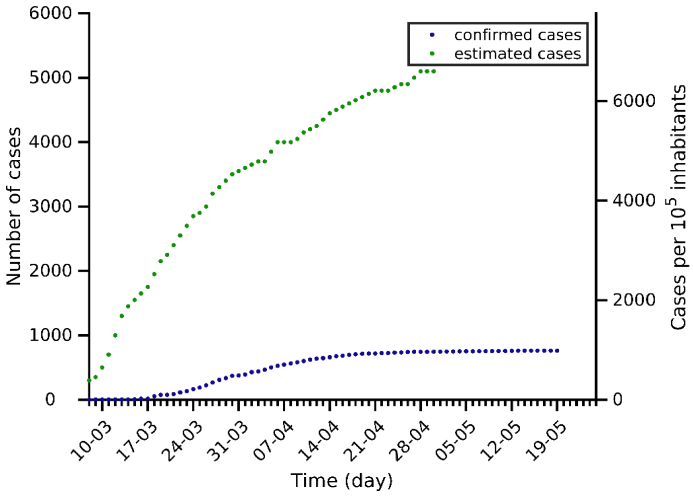


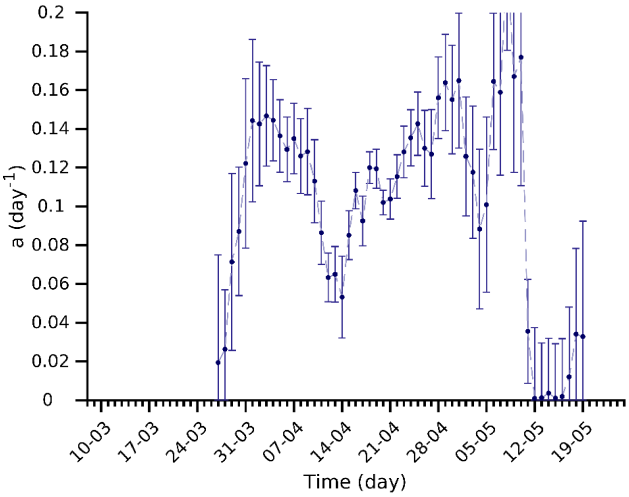


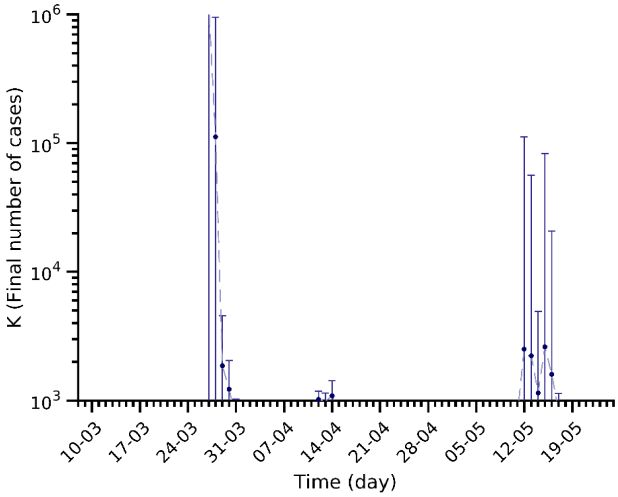


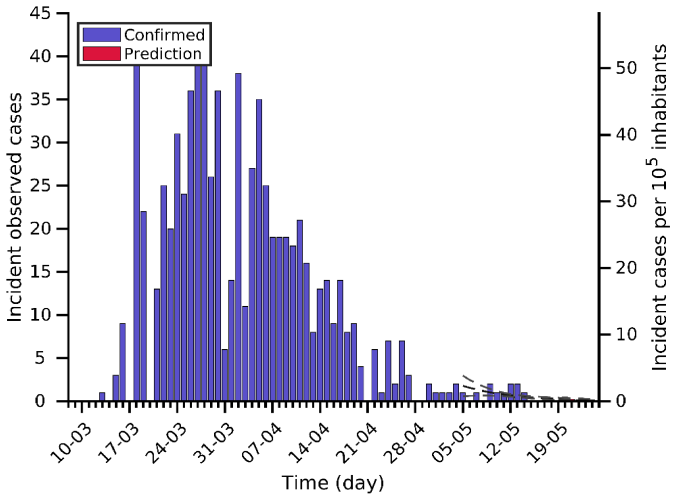
Andorra 19-05-2020. Population: 0.1M. Current cumulated incidence: 985/10^5

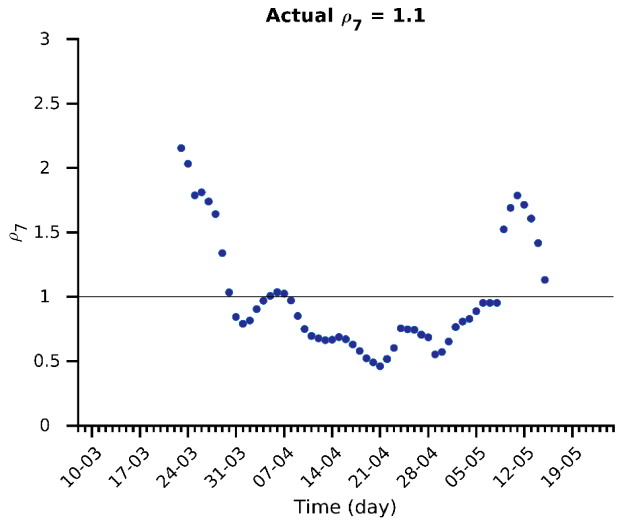


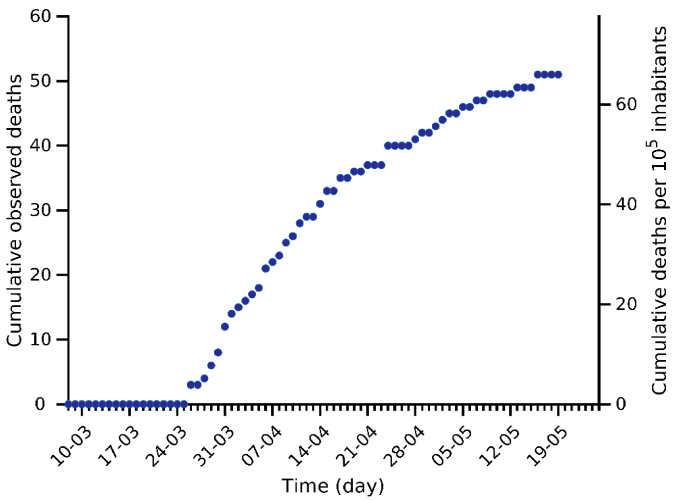


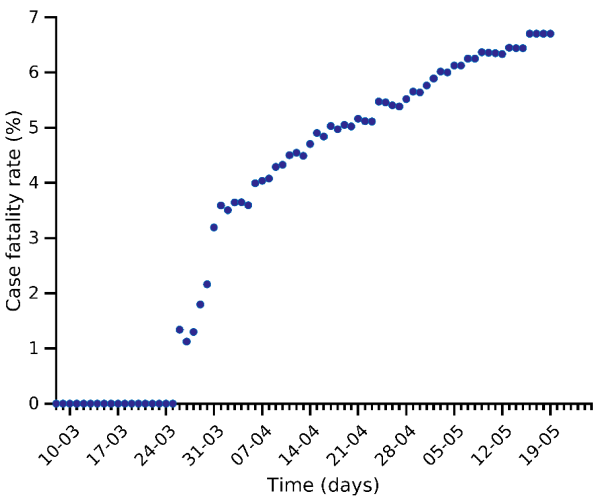








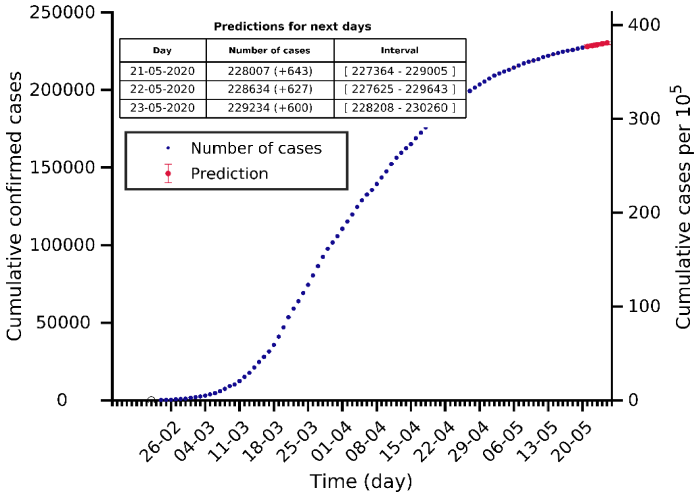


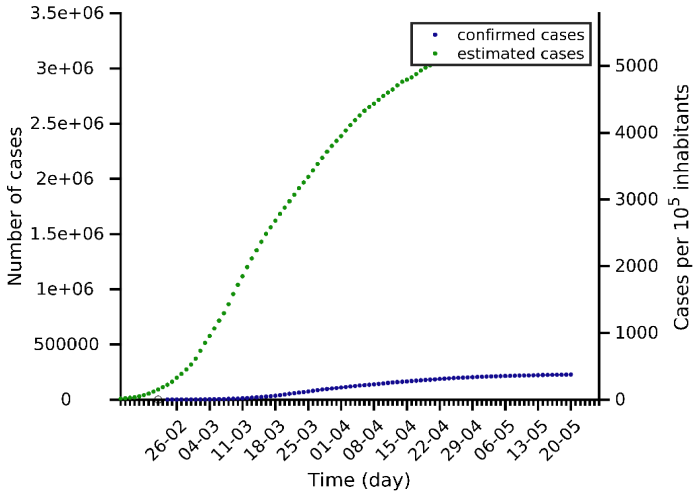


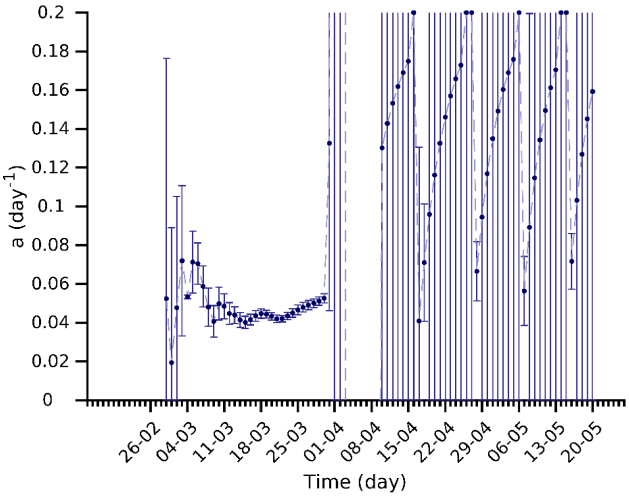
(3) Analysis and prediction of COVID-19 for Italy and its regions

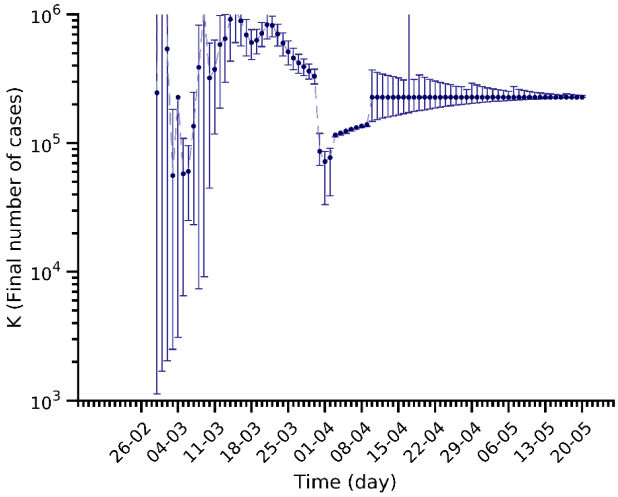
Data obtained from: https://github.com/pcm-dpc/COVID-19/tree/master/dati-andamento-nazionale

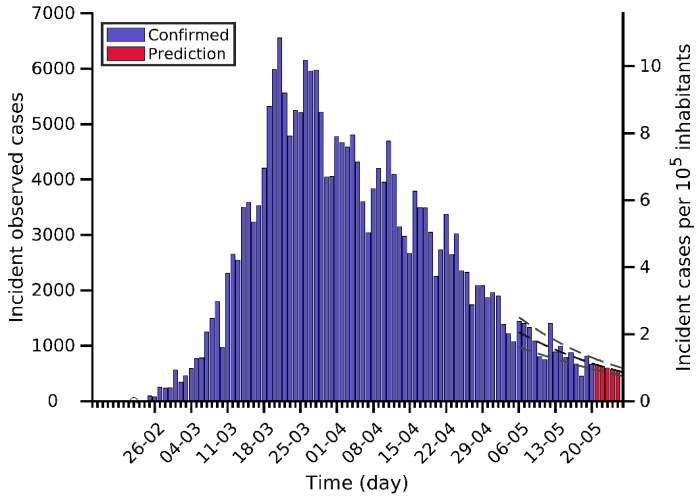
Italy 20-05-2020. Population: 60.5M. Current cumulated incidence: $376/10^5$

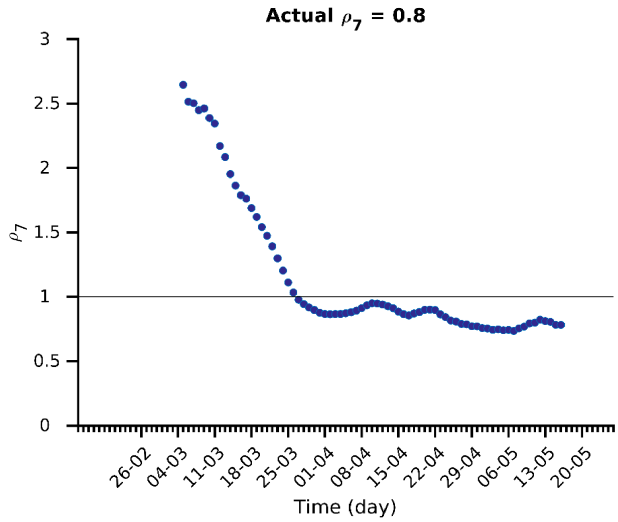


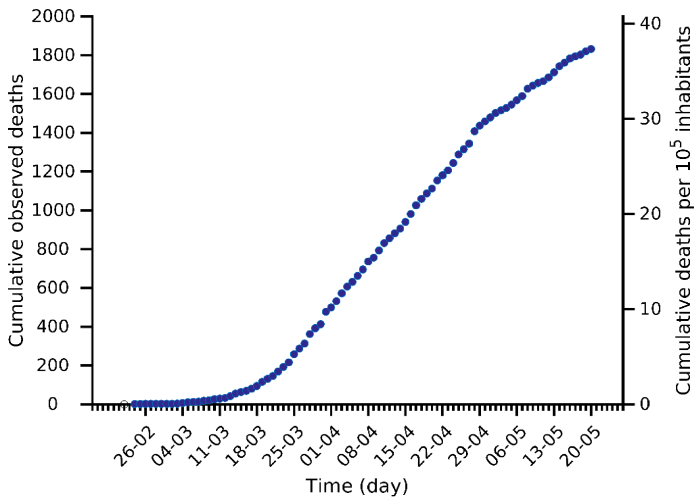


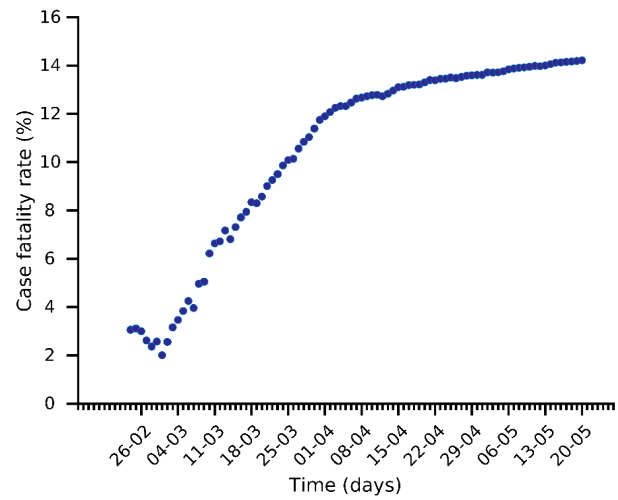




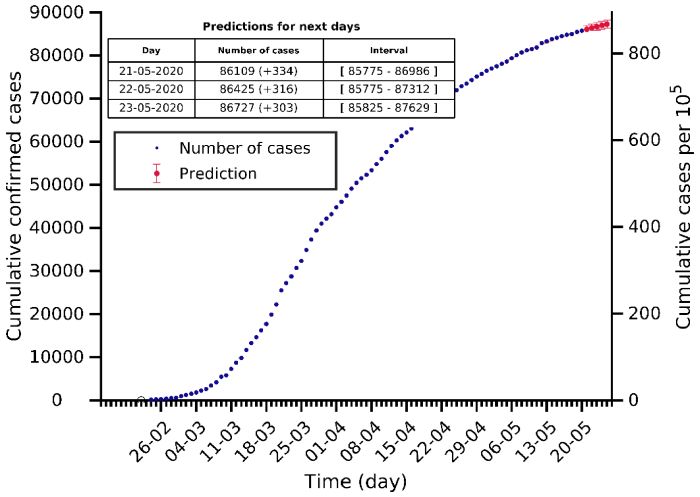


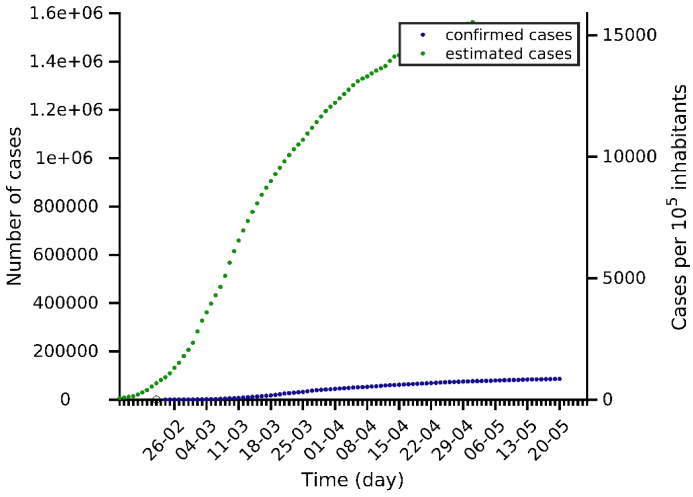


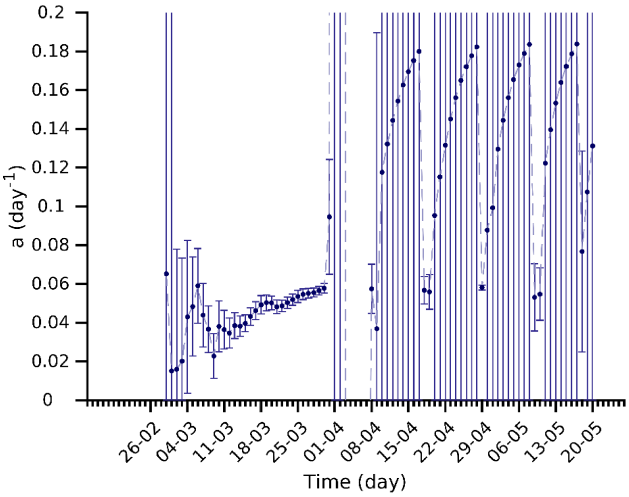


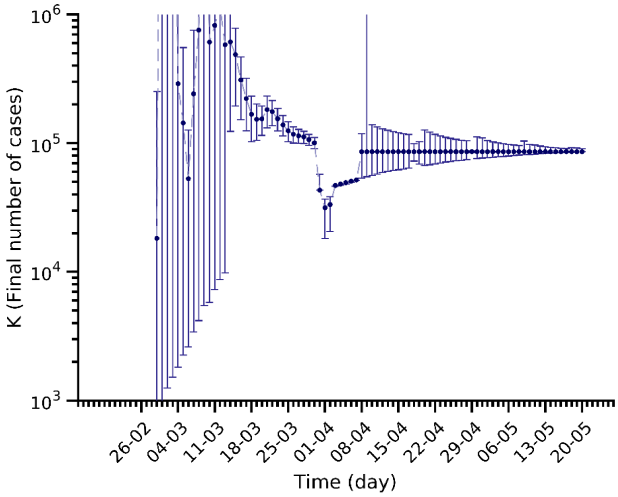


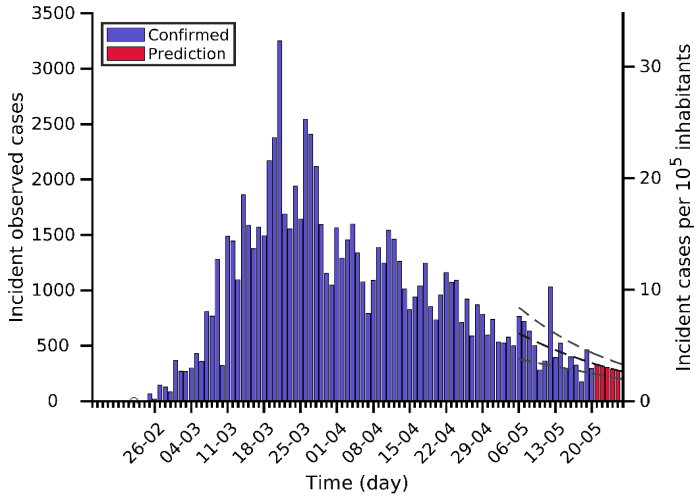
Lombardia 20-05-2020. Population: 10.1M. Current cumulated incidence: 853/10^5

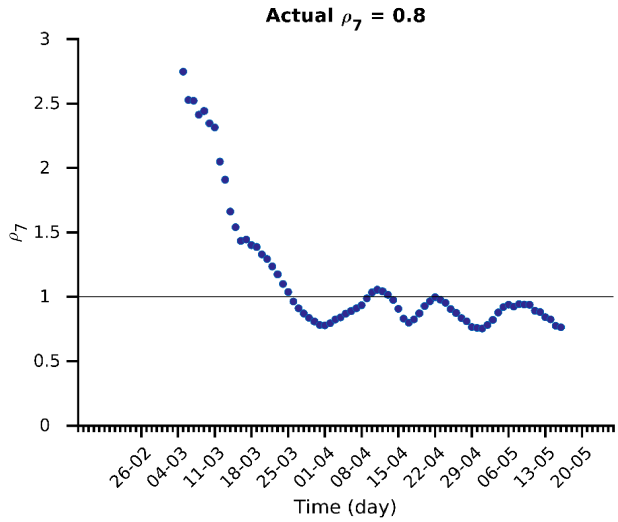


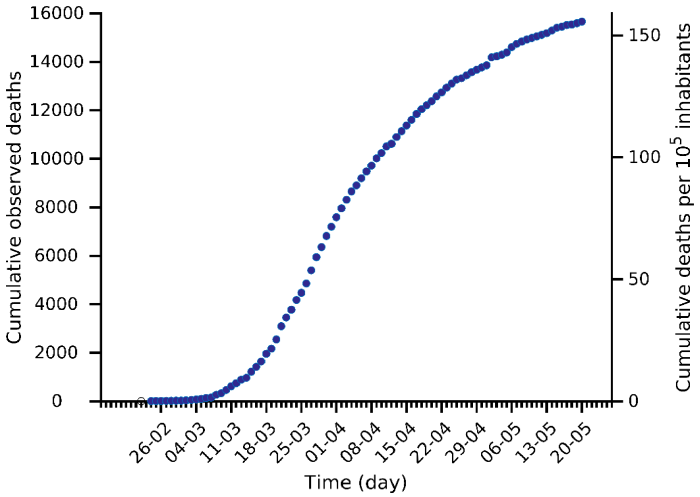


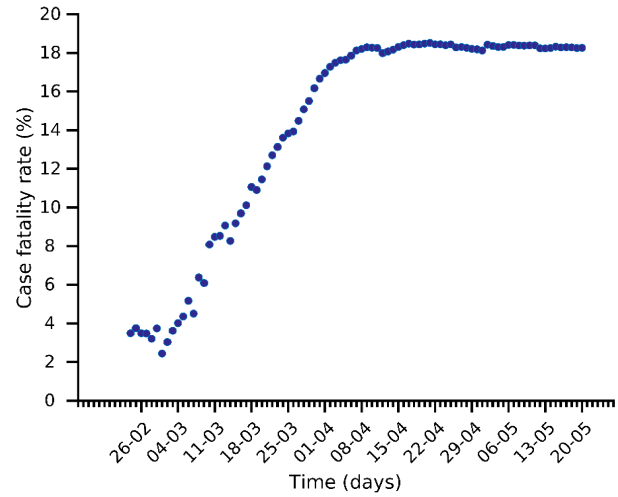




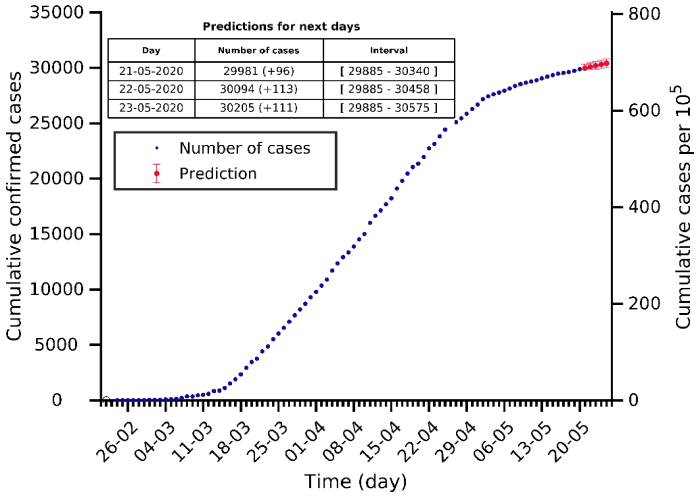


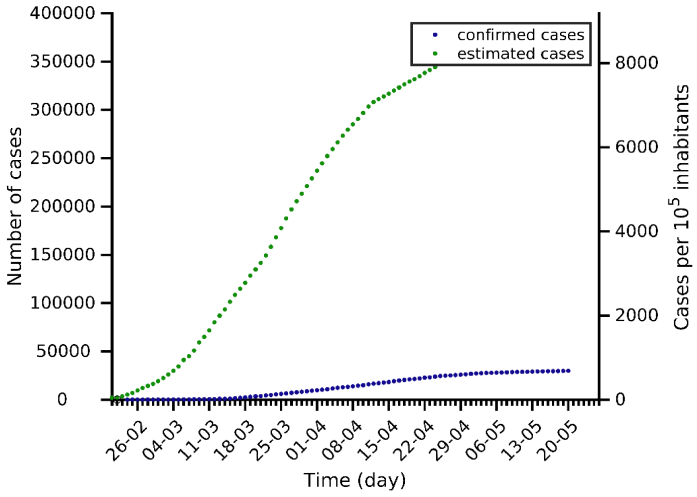


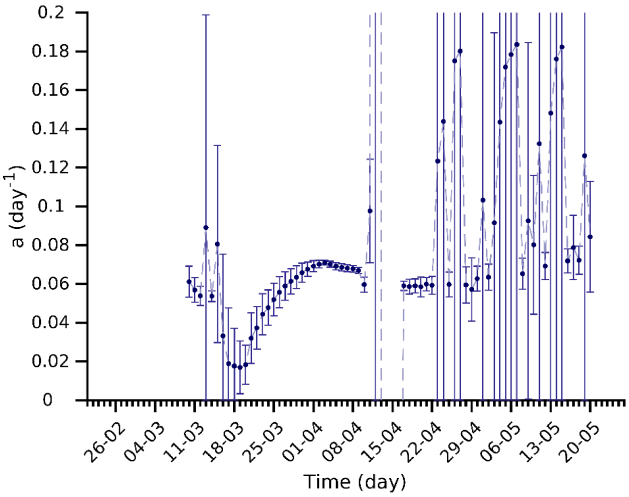


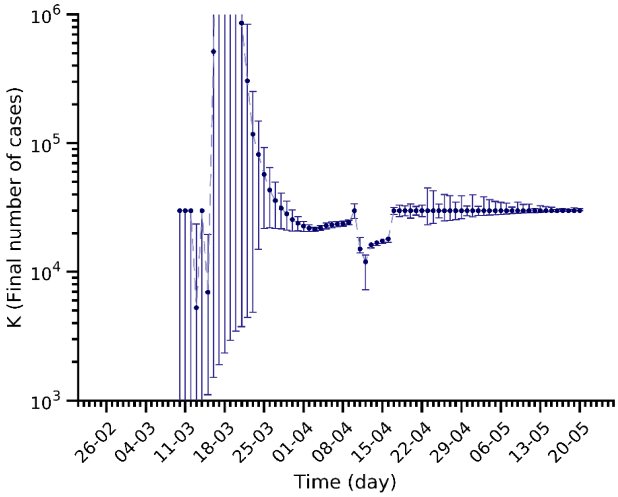


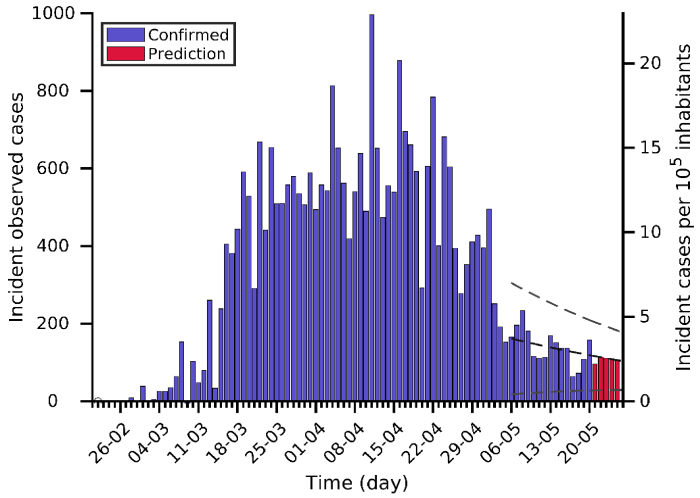
Piemonte 20-05-2020. Population: 4.4M. Current cumulated incidence: 686/10⁵

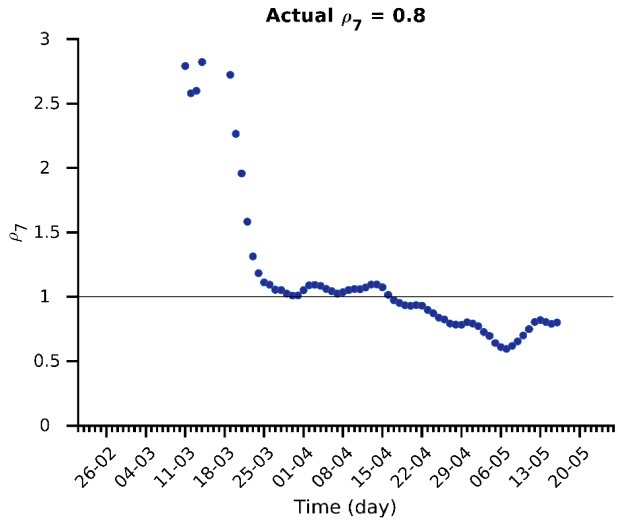


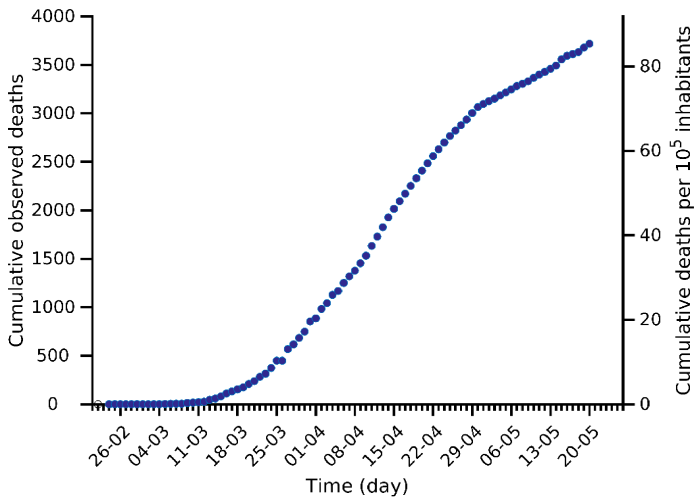


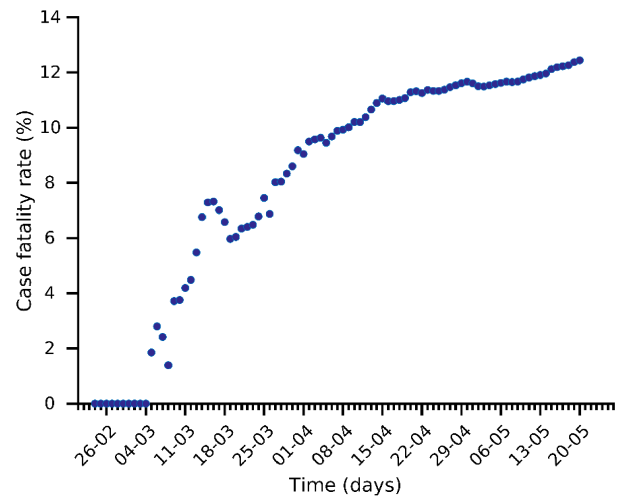




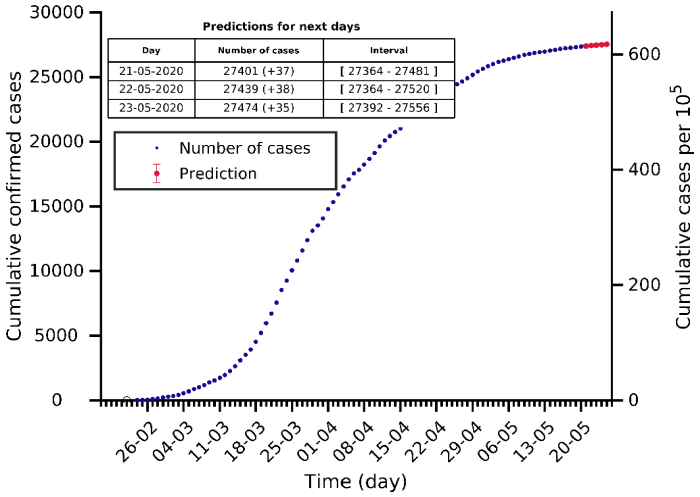


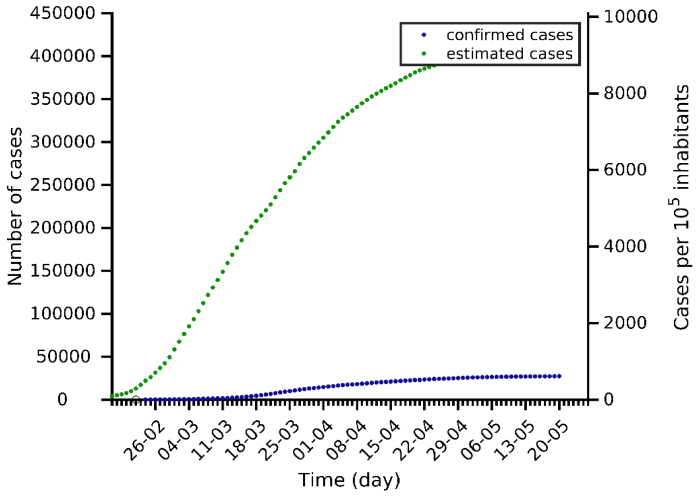


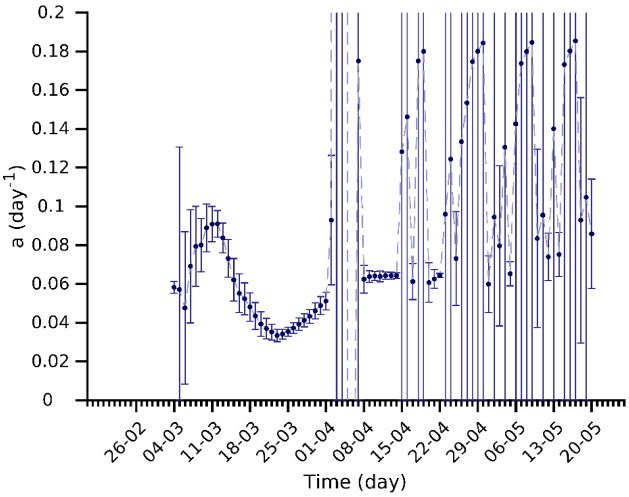


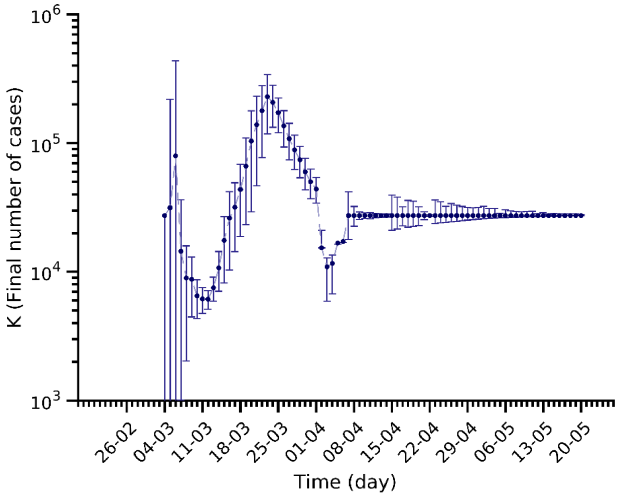


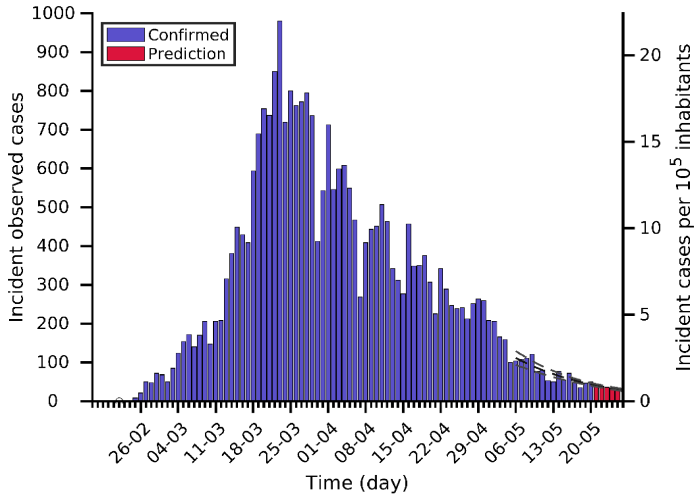
Emilia Romagna 20-05-2020. Population: 4.5M. Current cumulated incidence: 614/1

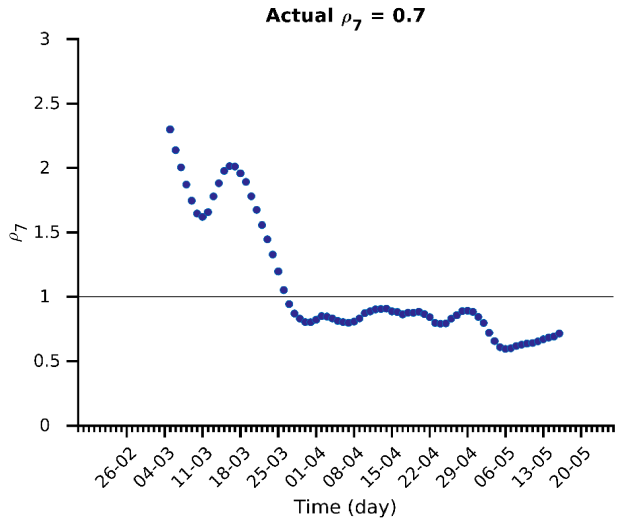


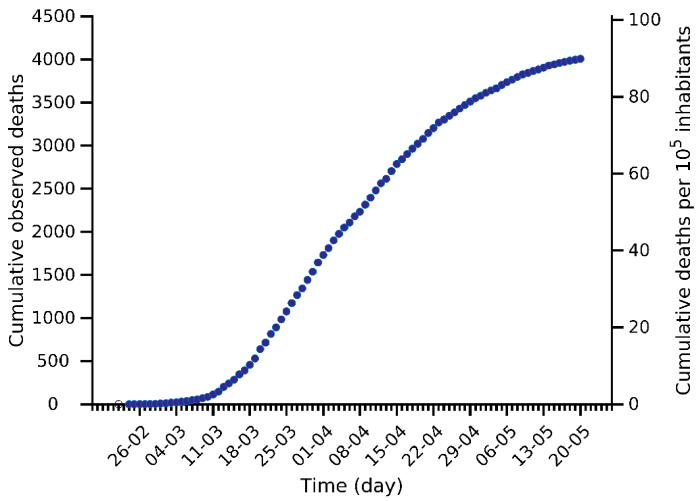


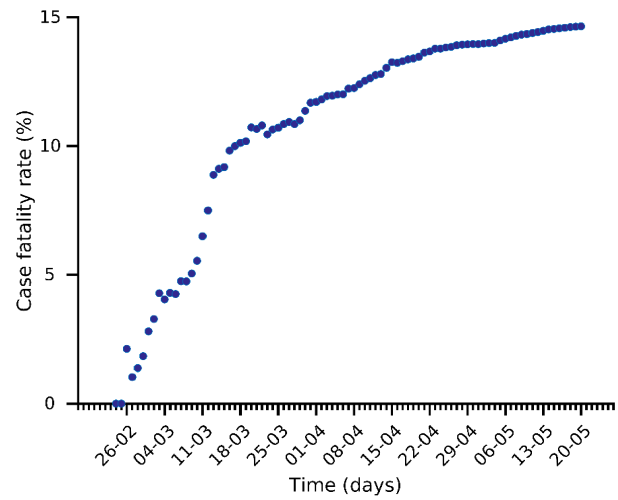




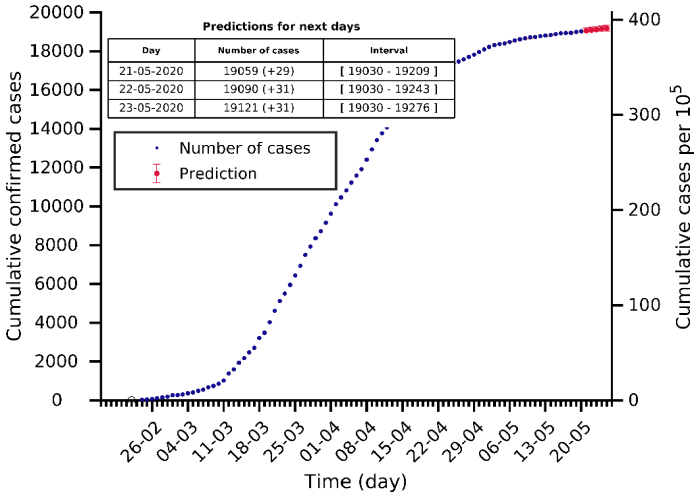


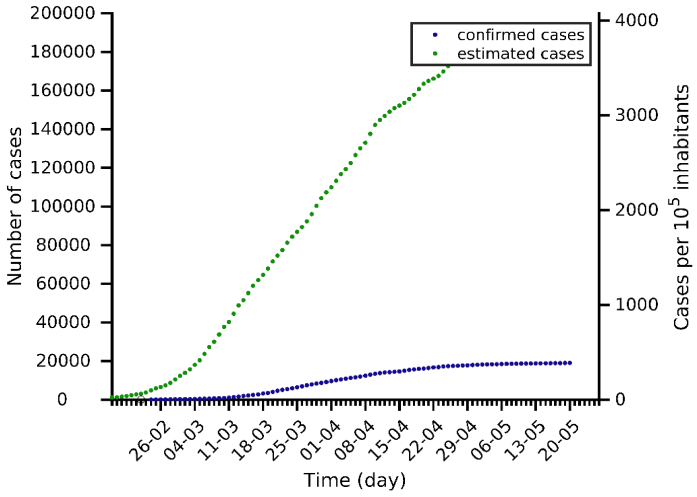


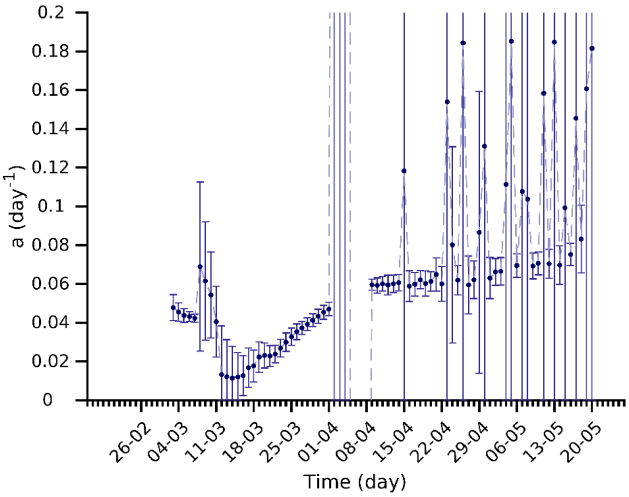


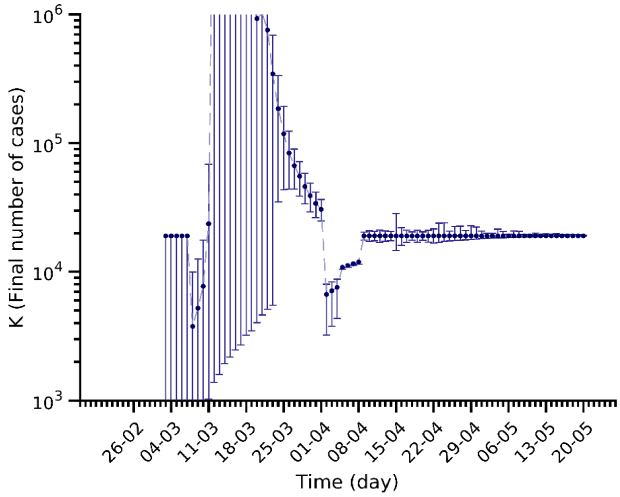


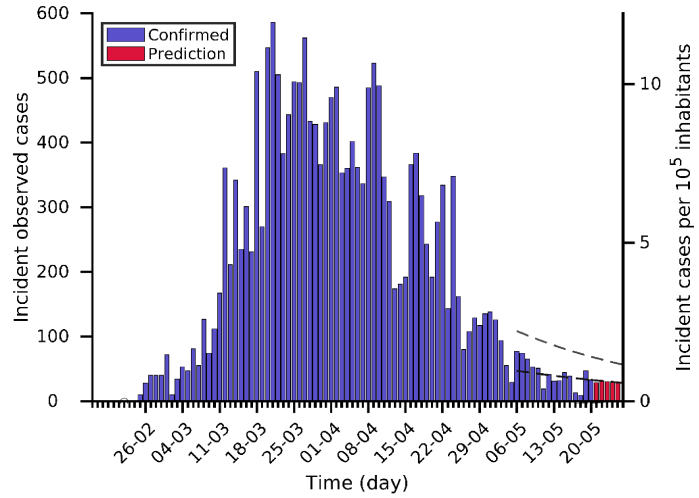
Veneto 20-05-2020. Population: 4.9M. Current cumulated incidence: 388/10^5

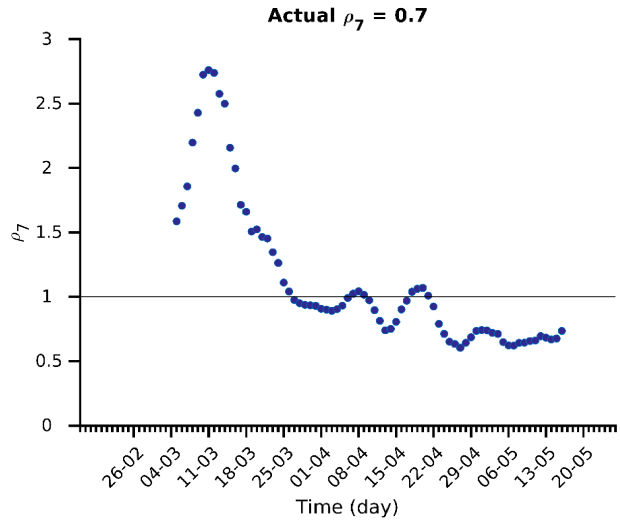


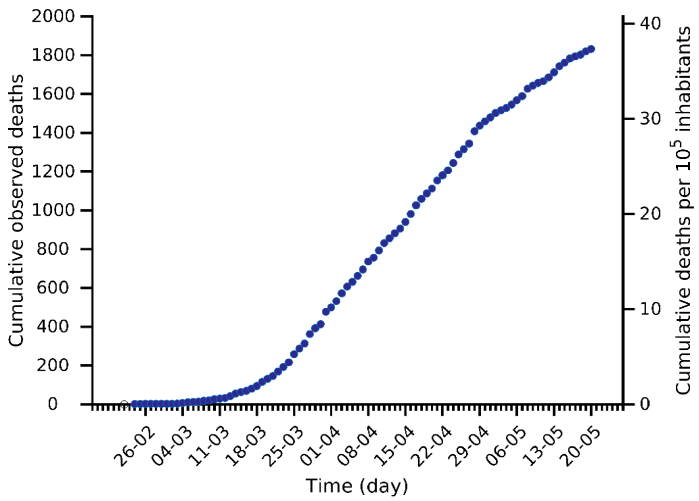


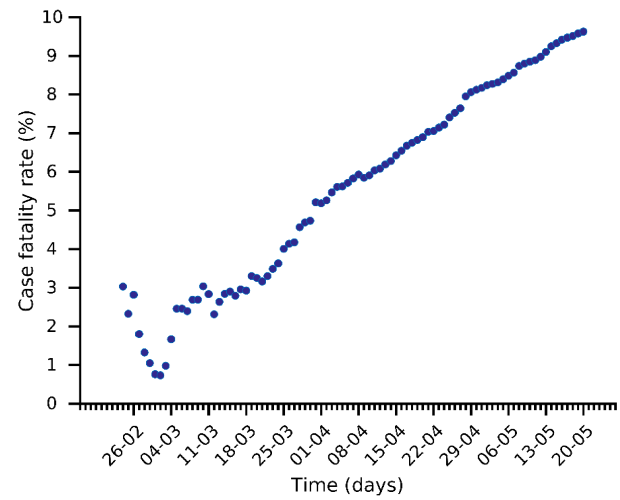




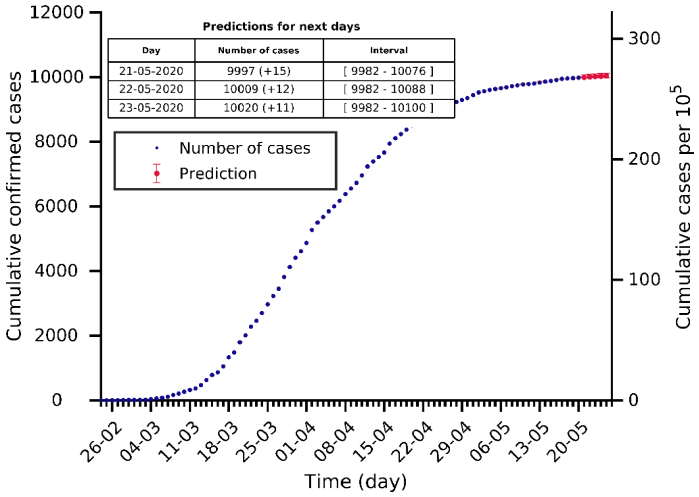


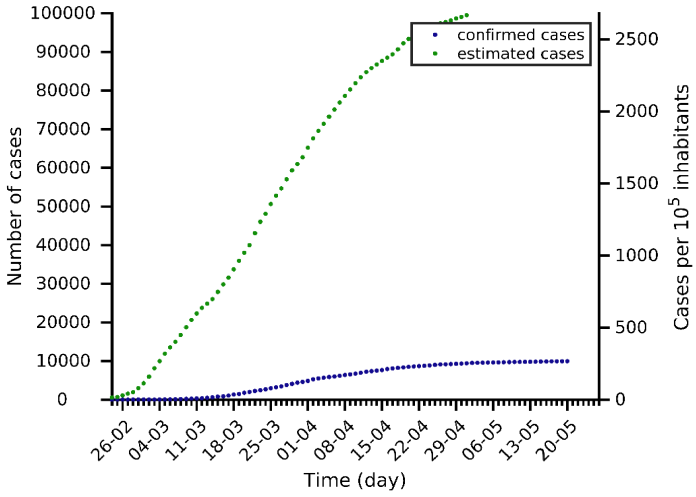


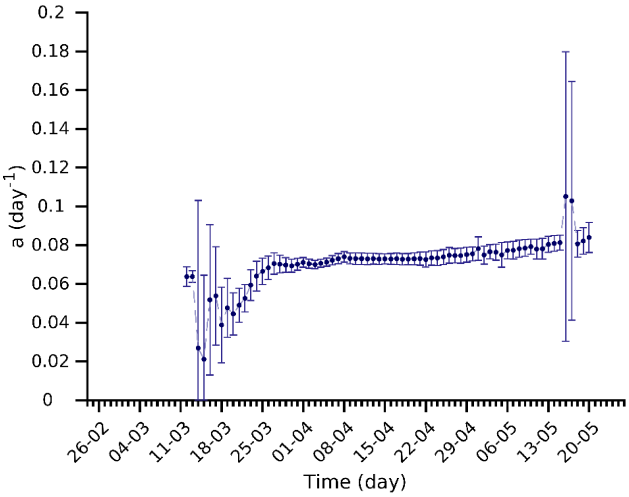


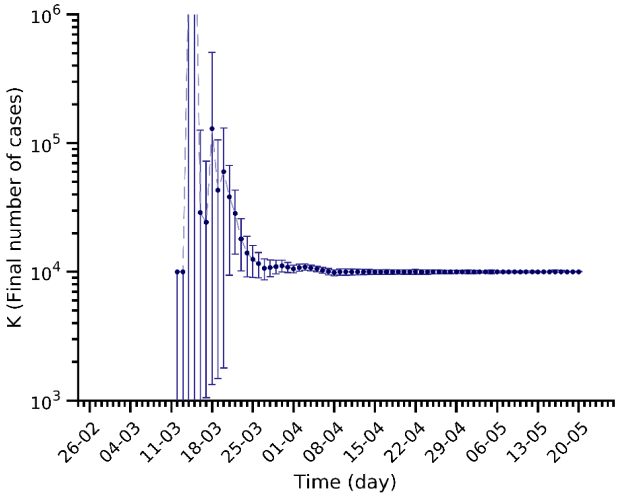


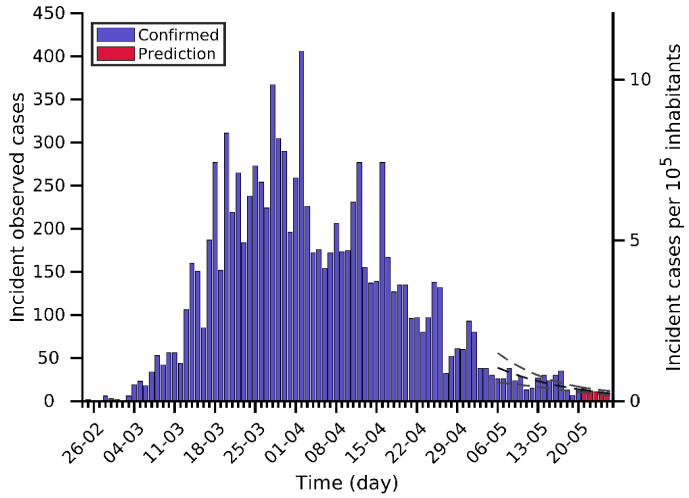
Toscana 20-05-2020. Population: 3.7M. Current cumulated incidence: 268/10⁵

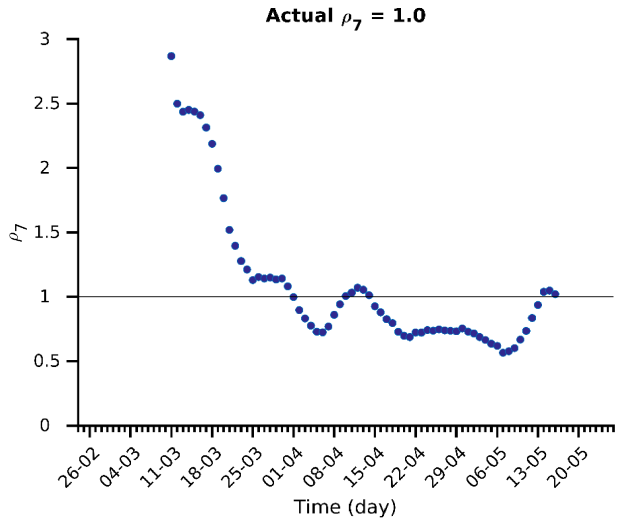


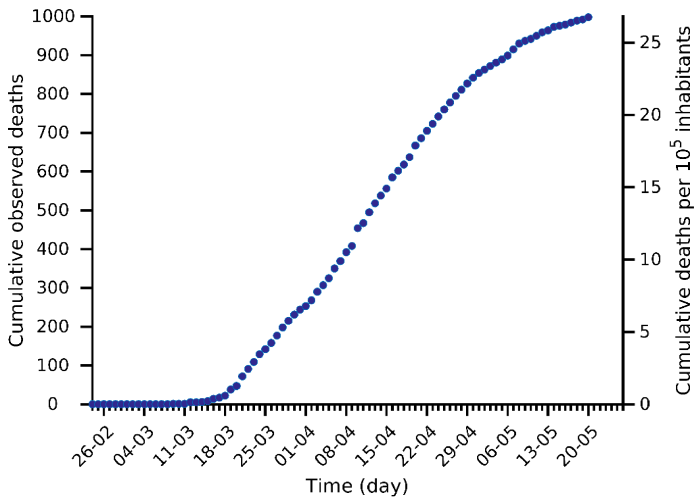


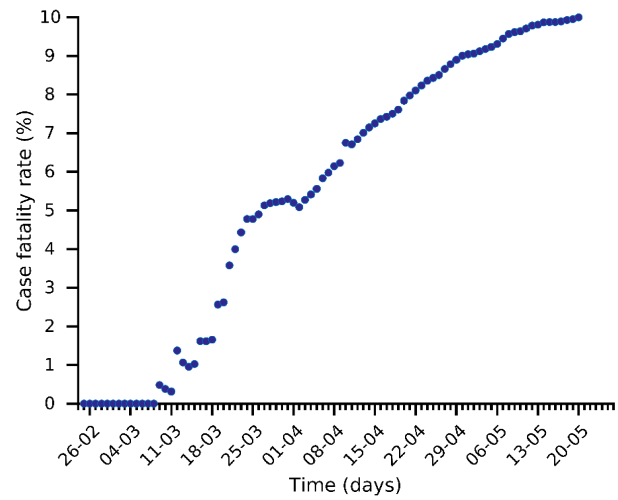




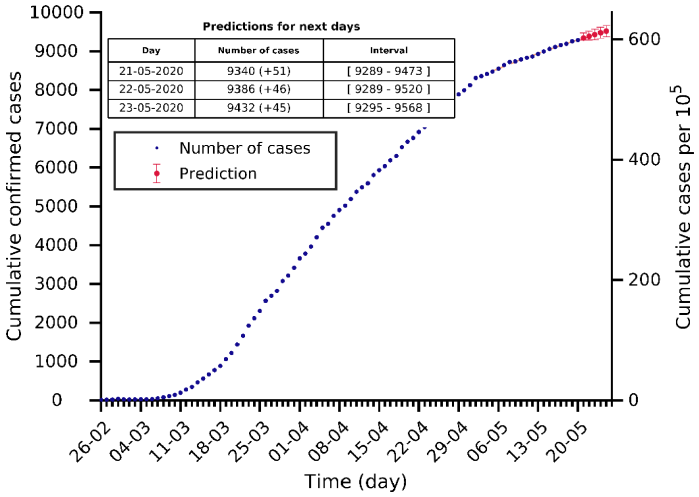


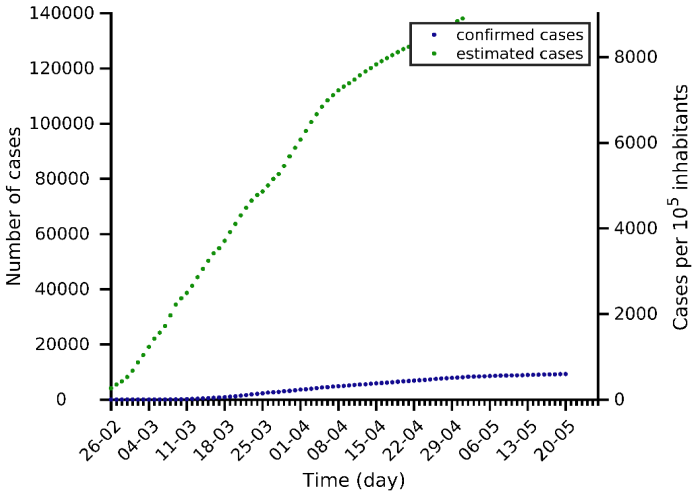


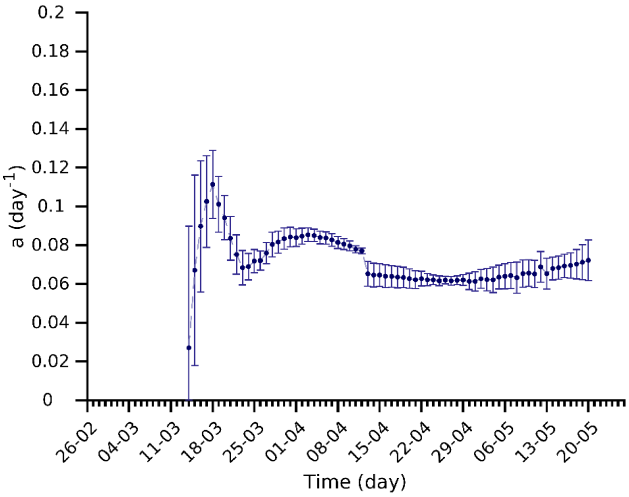


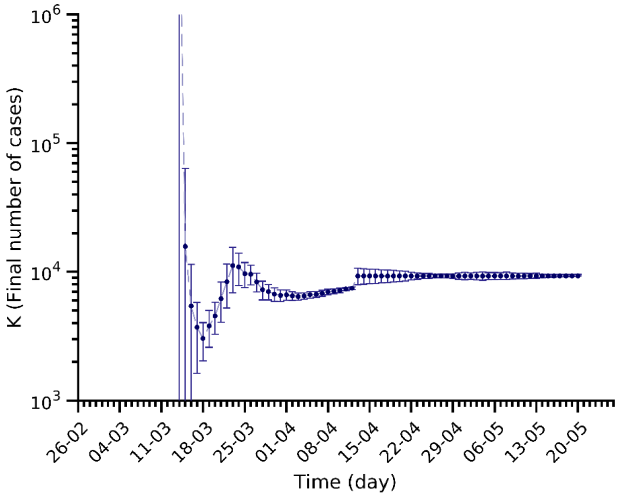


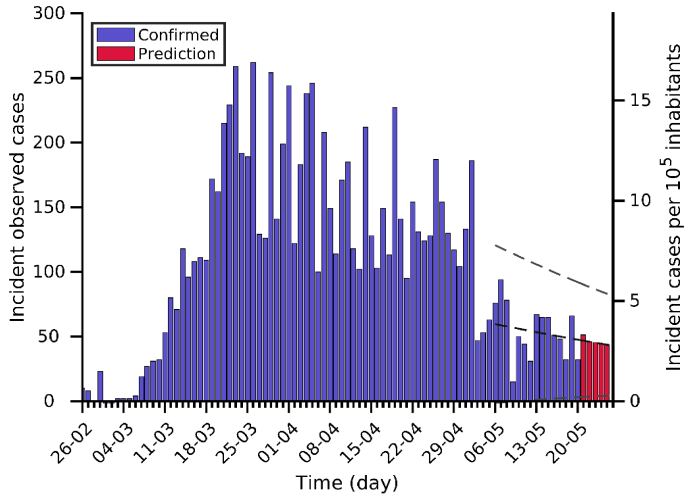
Liguria 20-05-2020. Population: 1.6M. Current cumulated incidence: 599/10^5

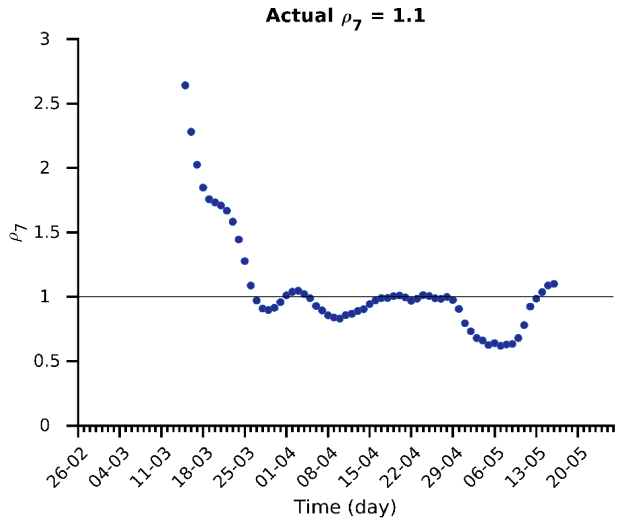


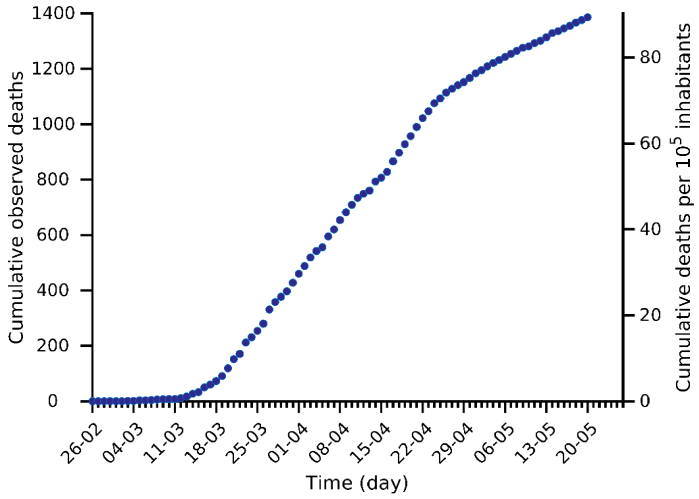


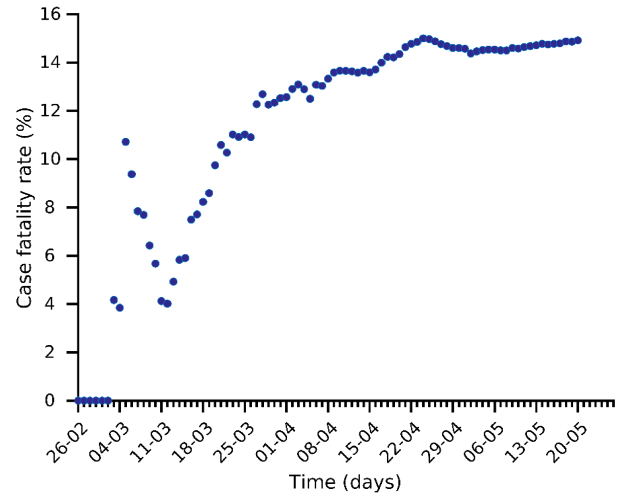




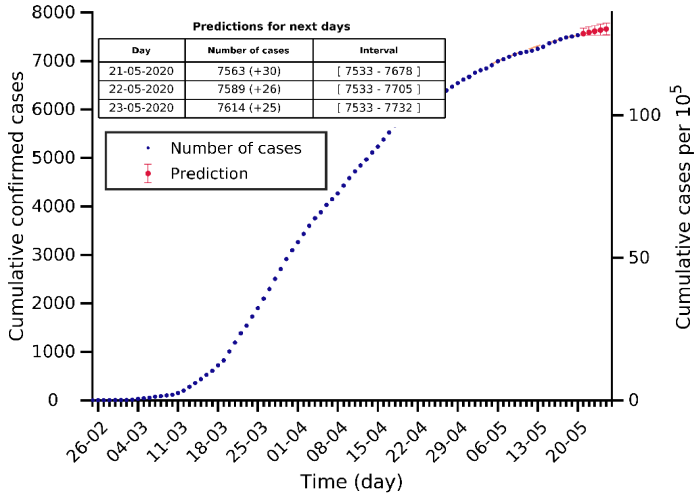


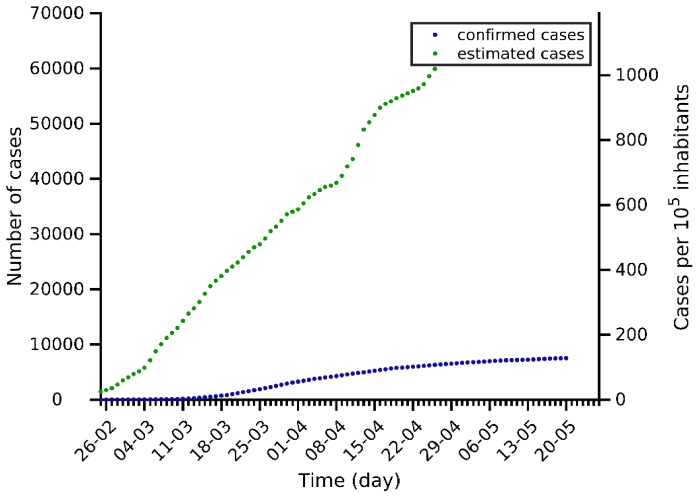


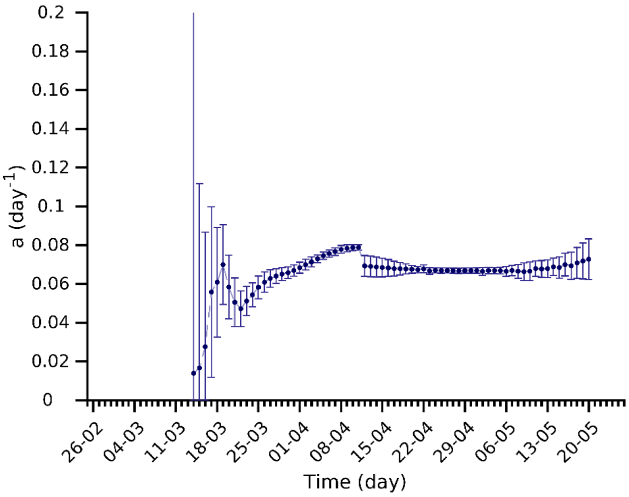


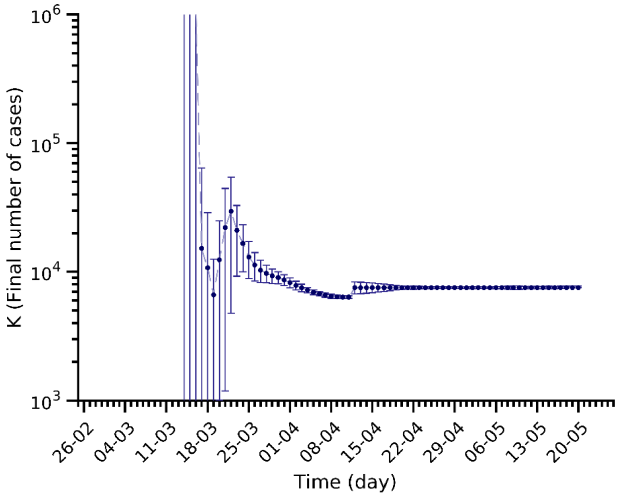


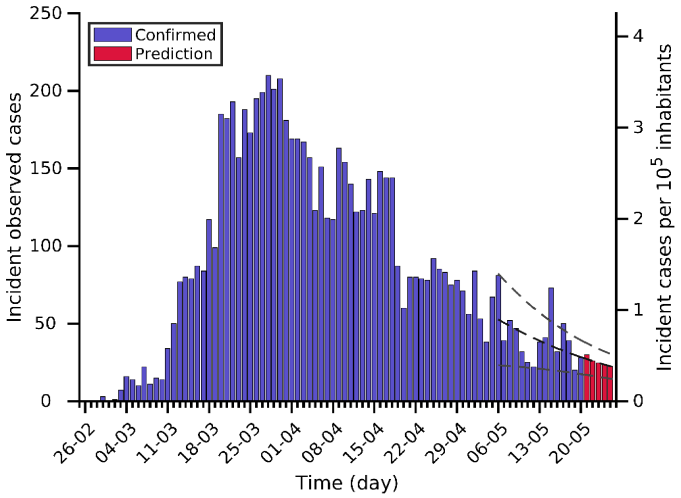
Lazio 20-05-2020. Population: 5.9M. Current cumulated incidence: 128/10⁵

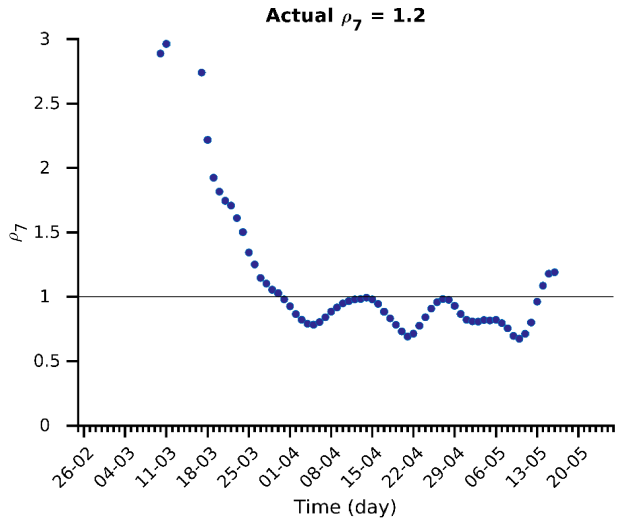


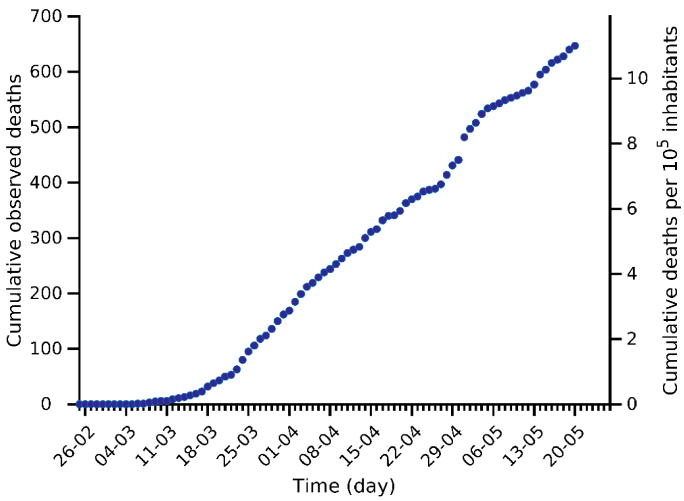


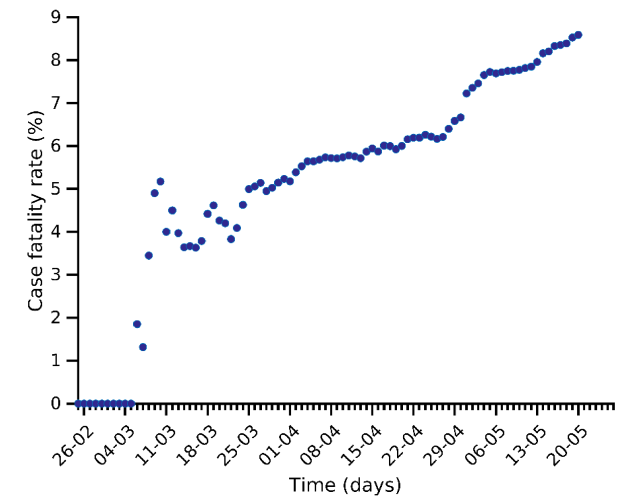




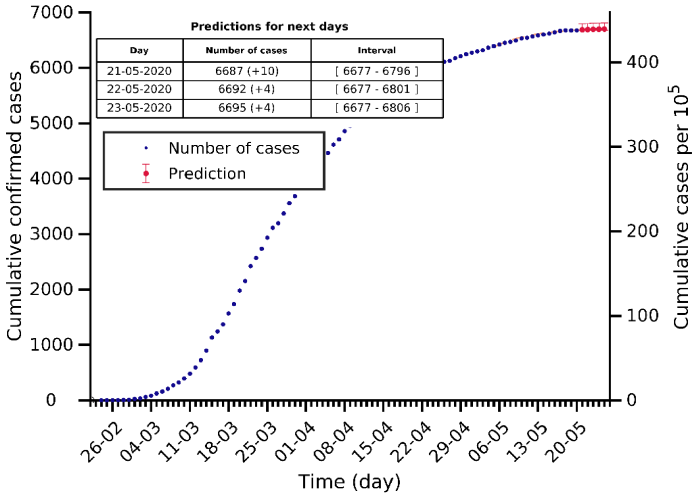


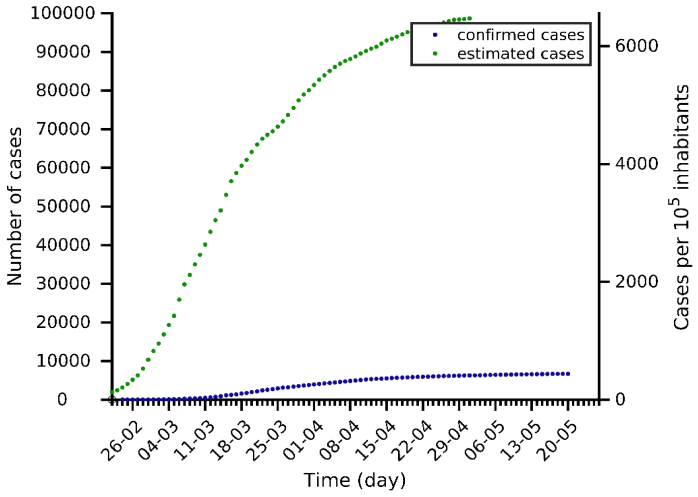


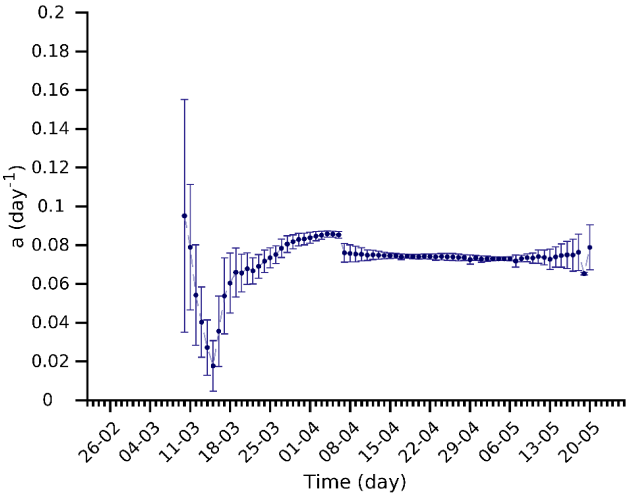


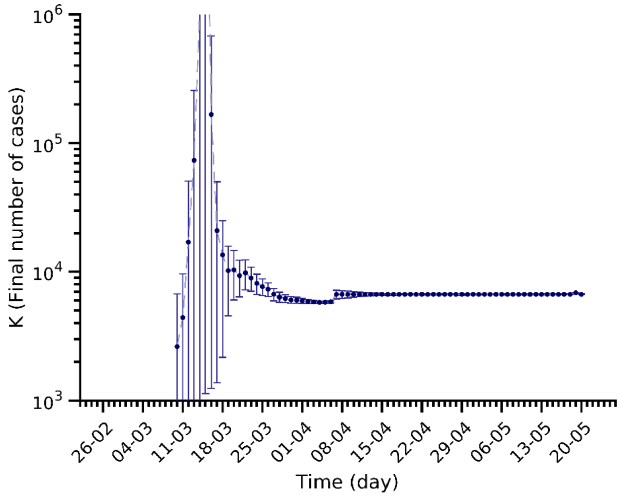


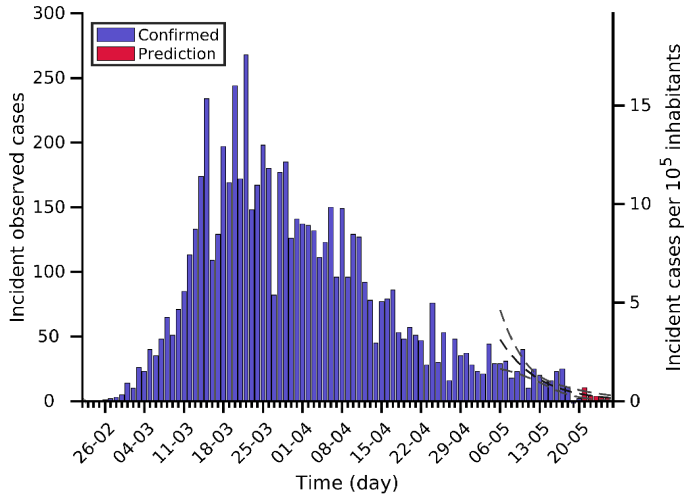
Marche 20-05-2020. Population: 1.5M. Current cumulated incidence: $438/10^5$

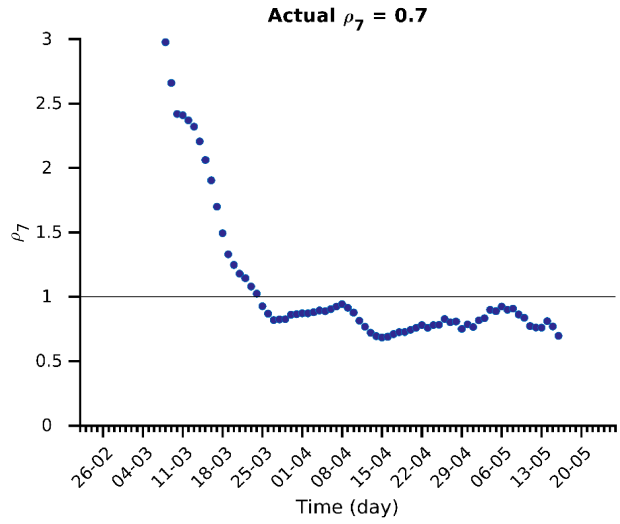


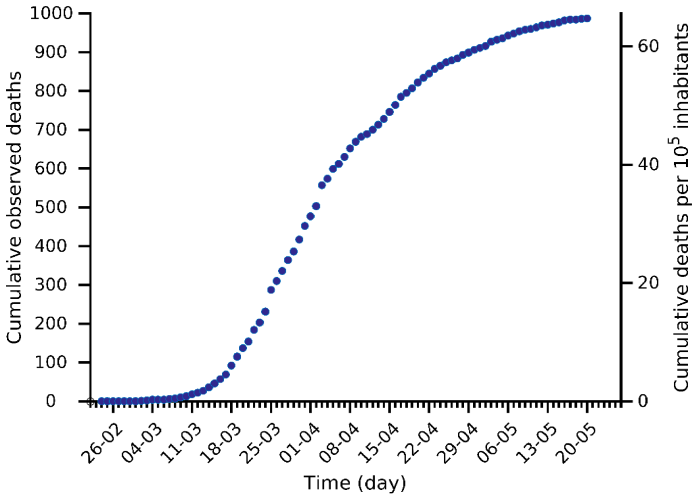


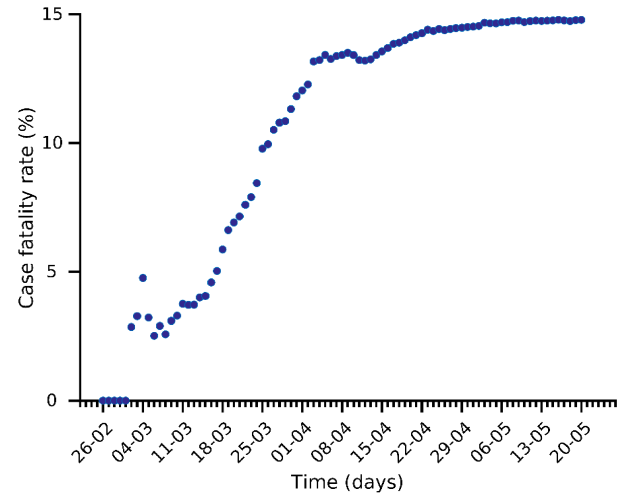




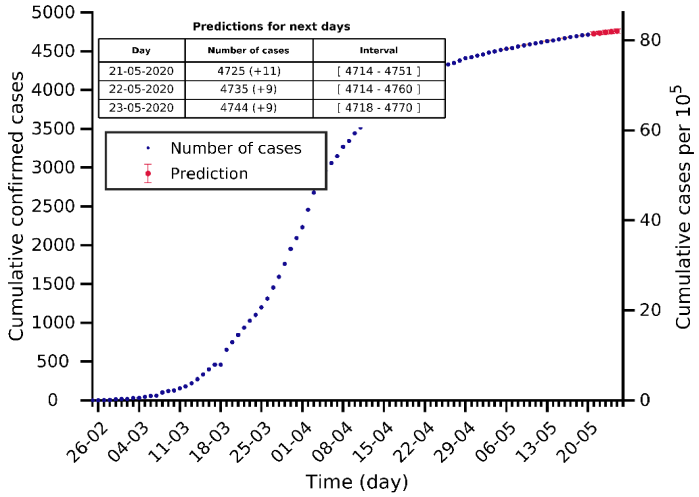


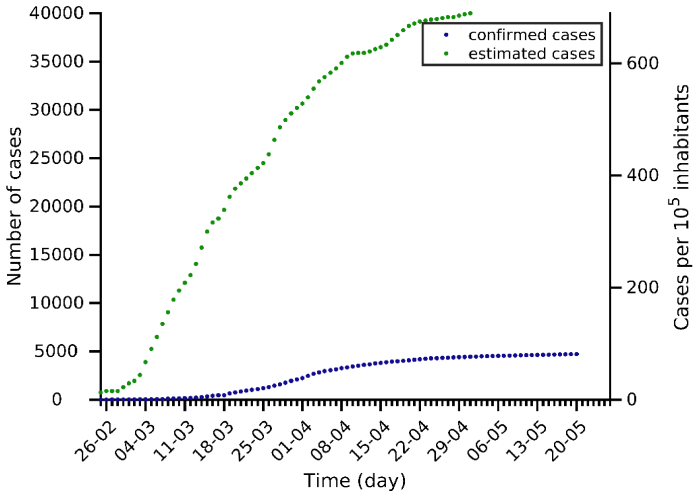


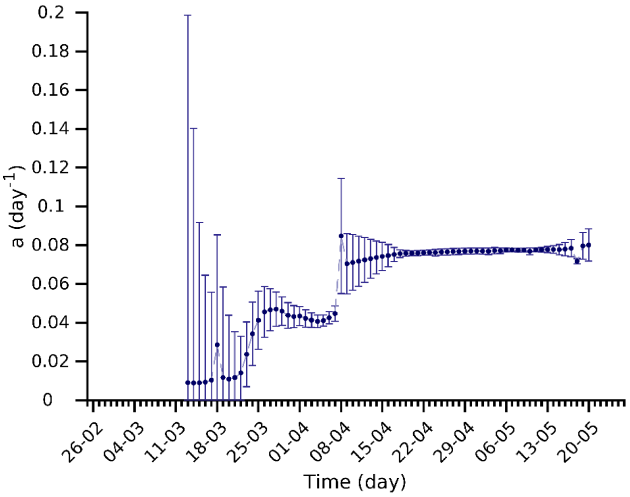


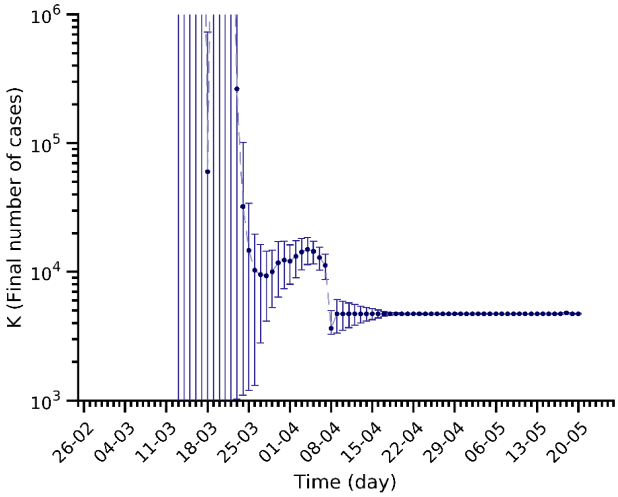


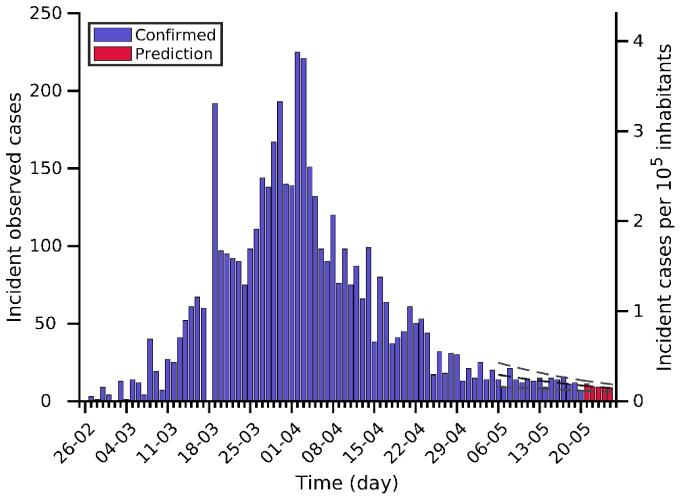
Campania 20-05-2020. Population: 5.8M. Current cumulated incidence: 81/10^5

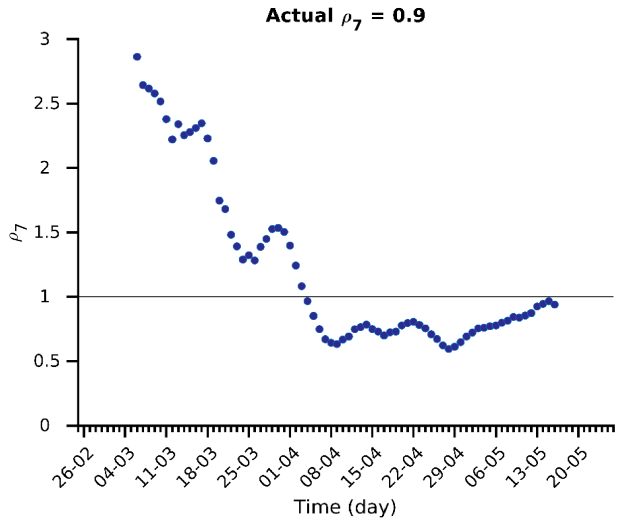


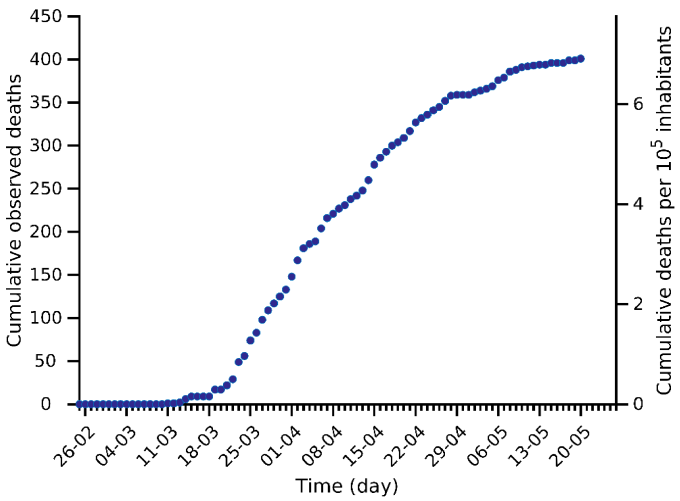


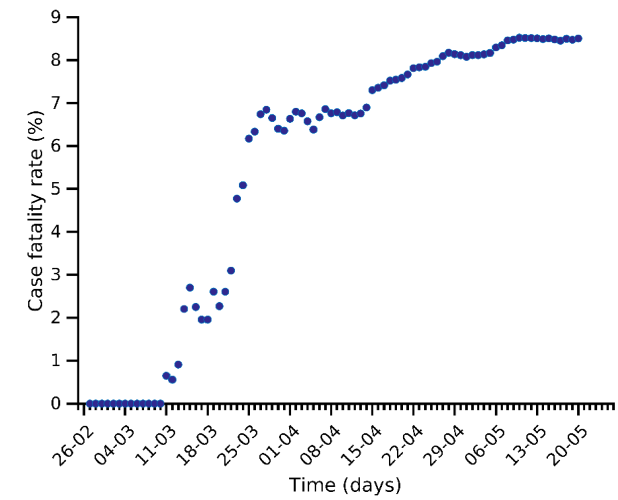




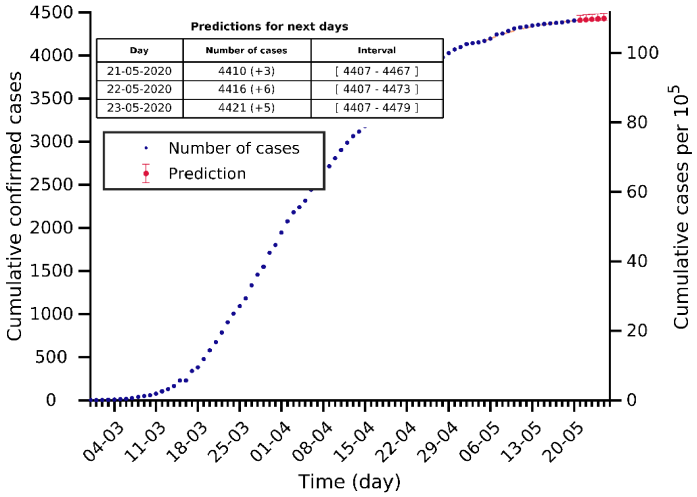


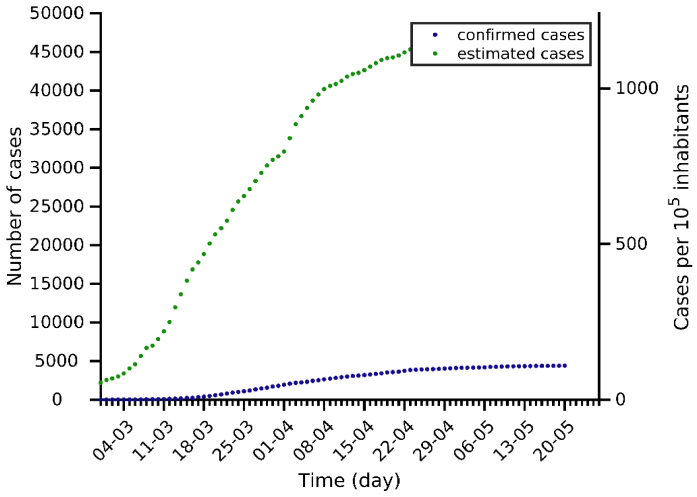


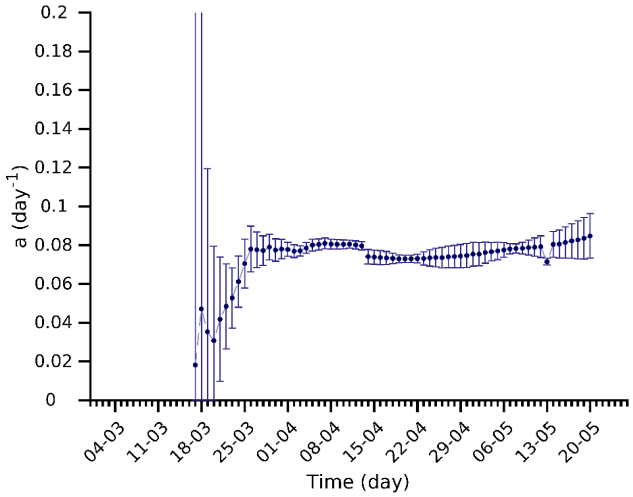


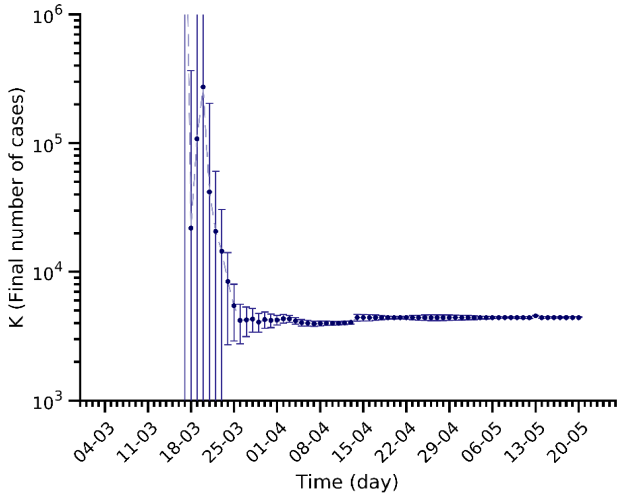


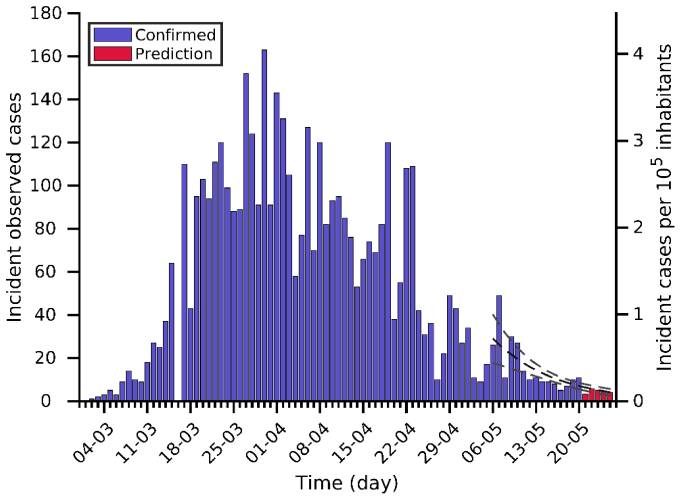
Puglia 20-05-2020. Population: 4.0M. Current cumulated incidence: $109/10^5$

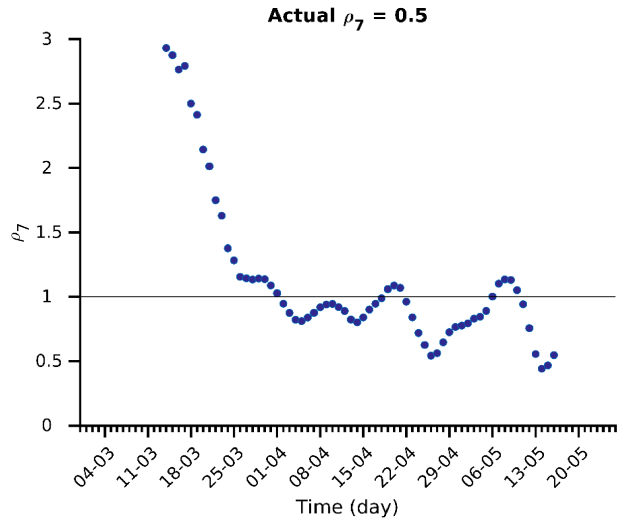


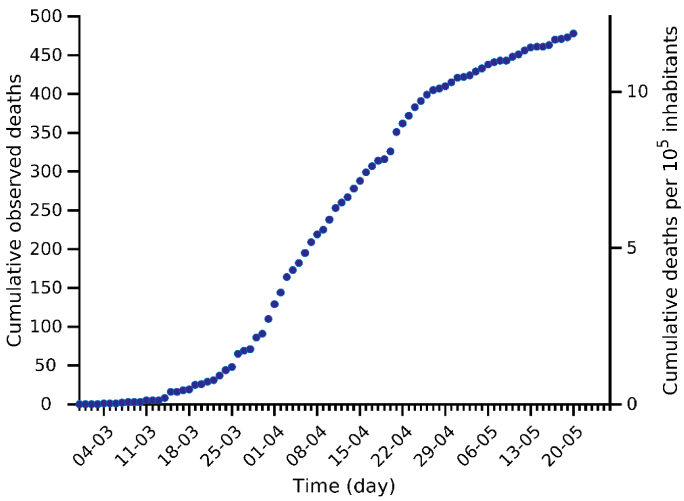


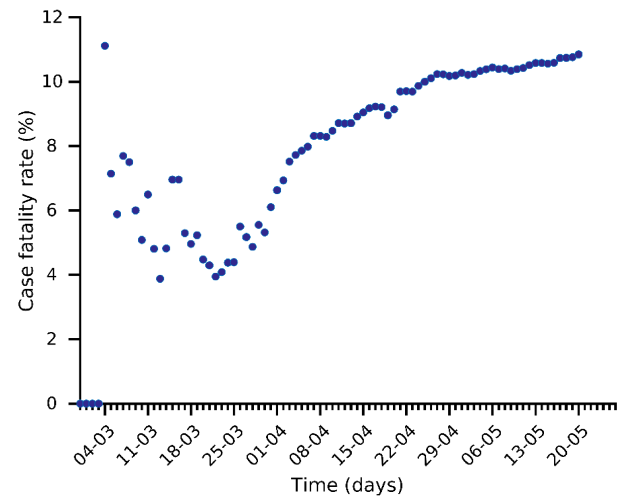




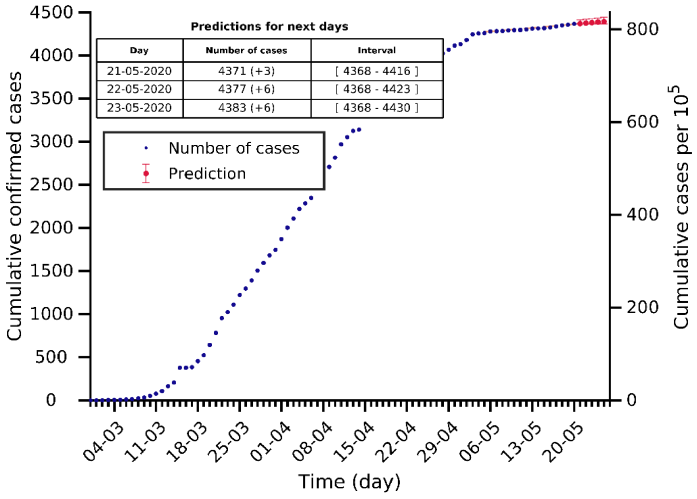


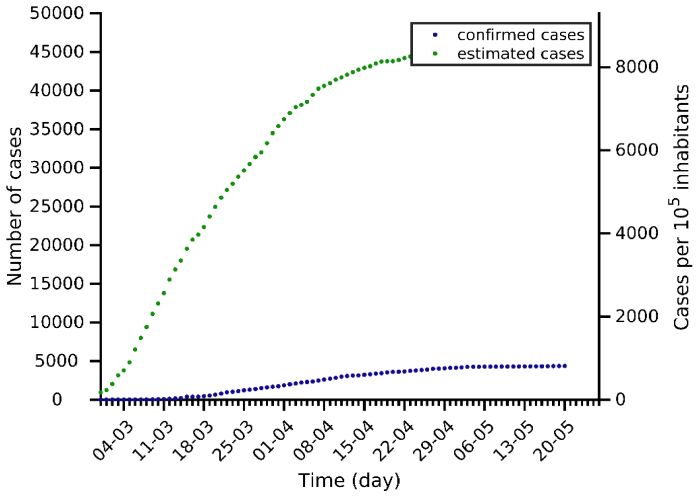


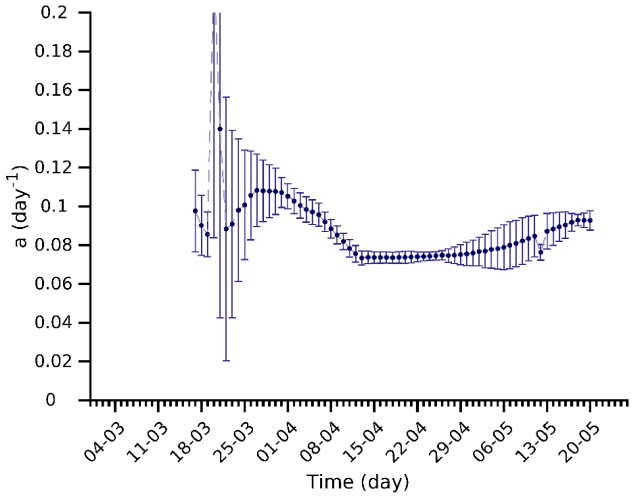


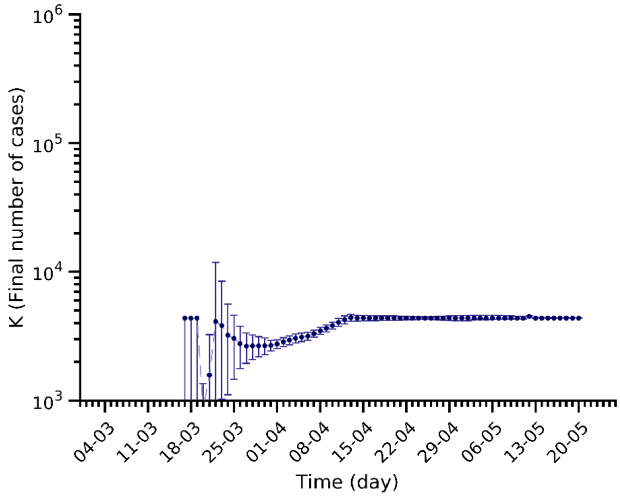


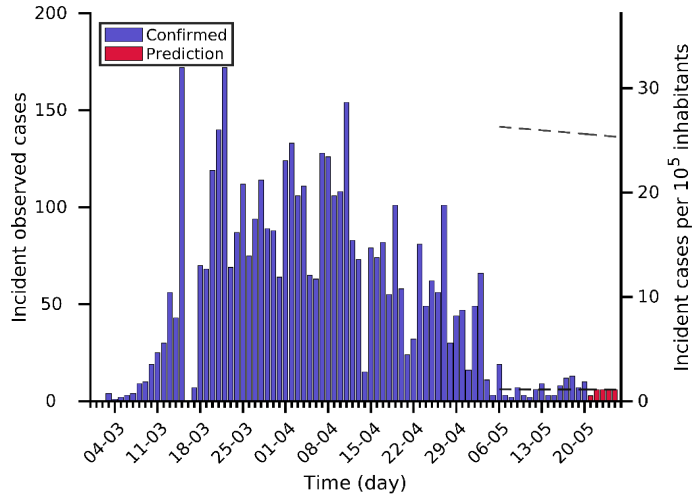
Trento 20-05-2020. Population: 0.5M. Current cumulated incidence: 812/10⁵

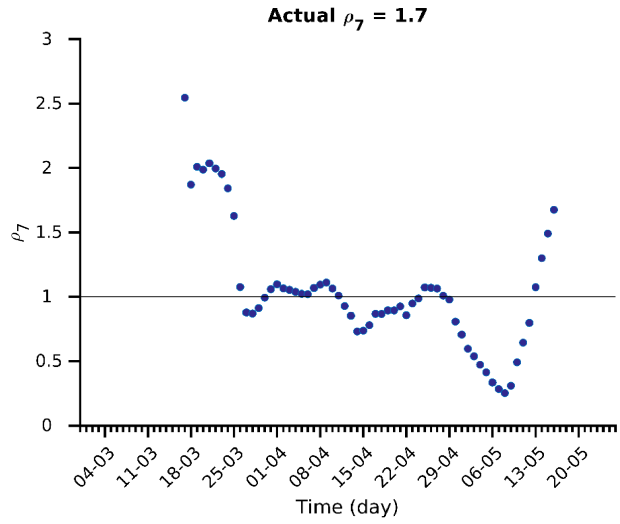


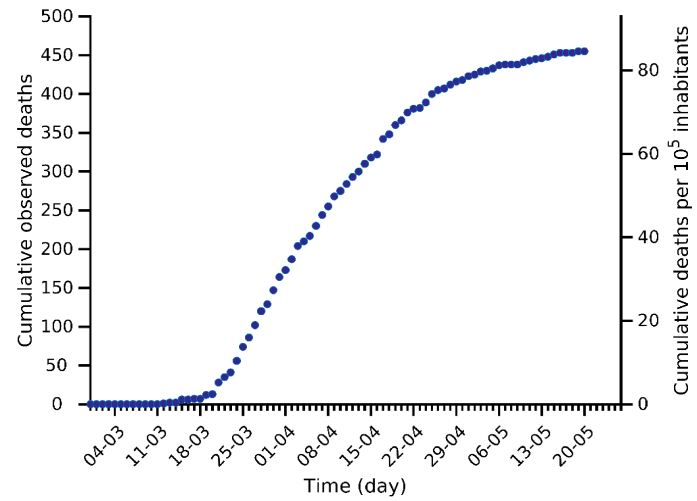


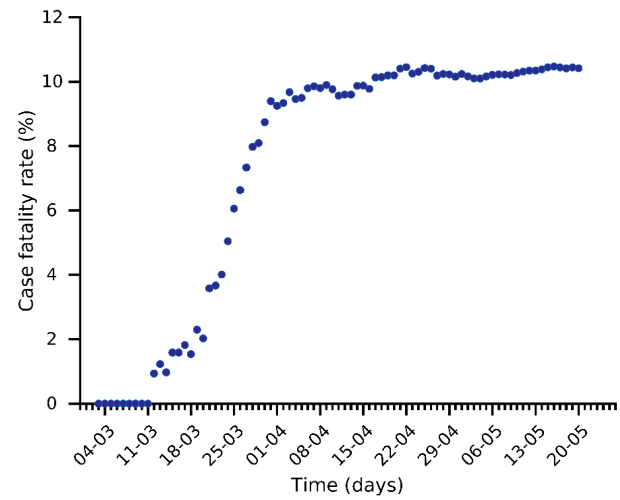




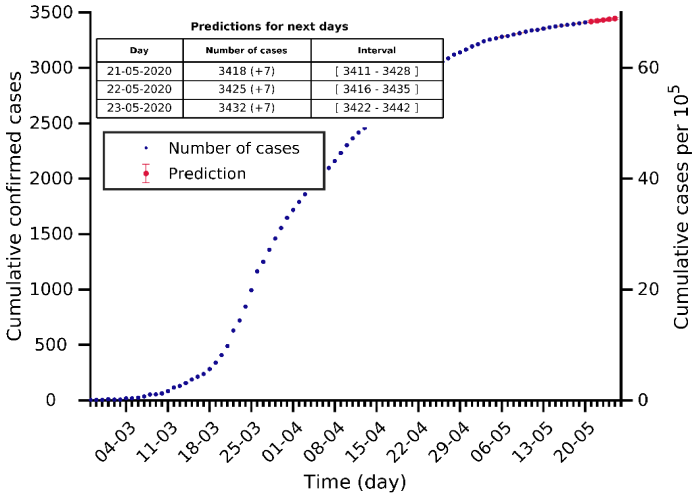


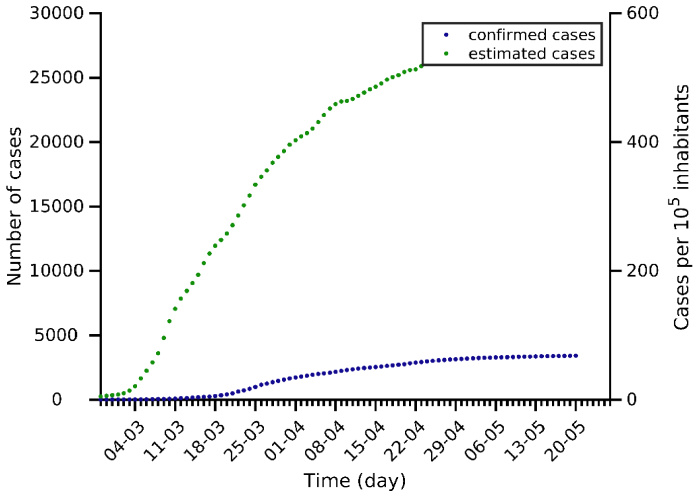


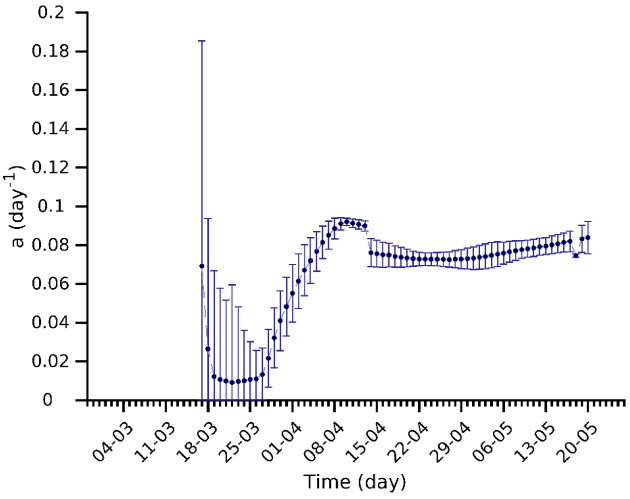


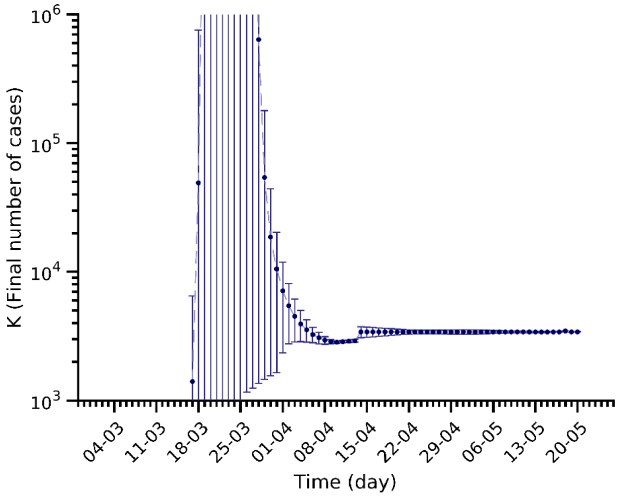


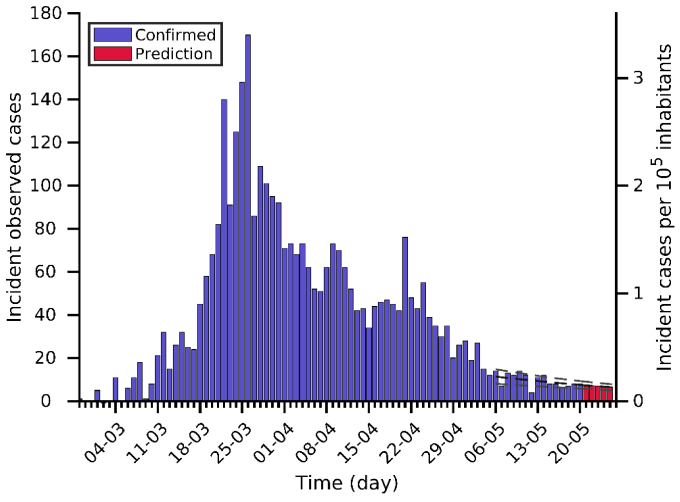
Sicilia 20-05-2020. Population: 5.0M. Current cumulated incidence: 68/10^5

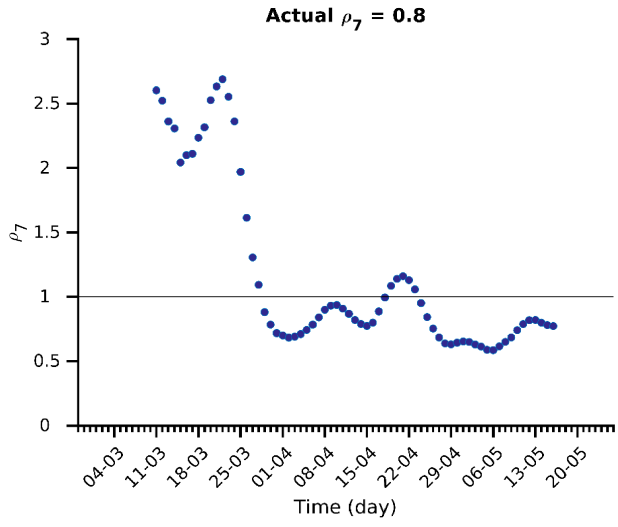


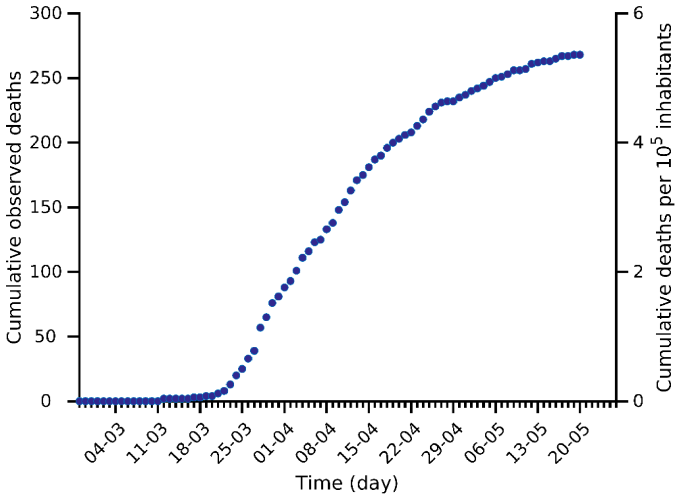


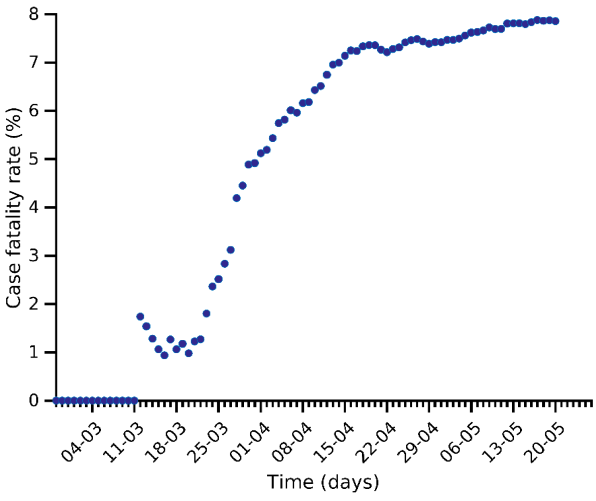




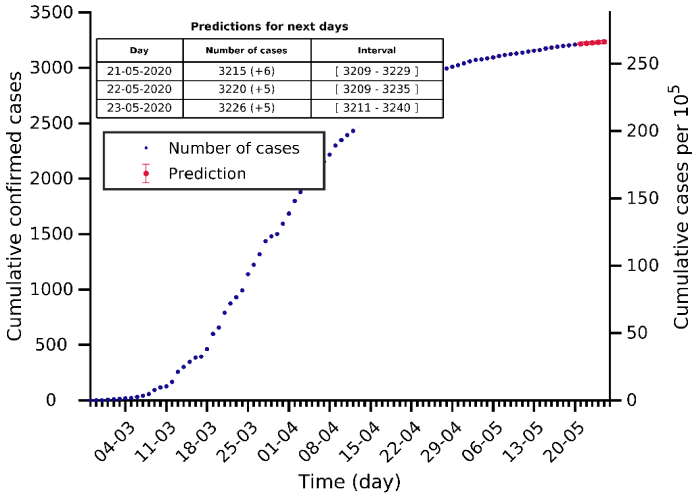


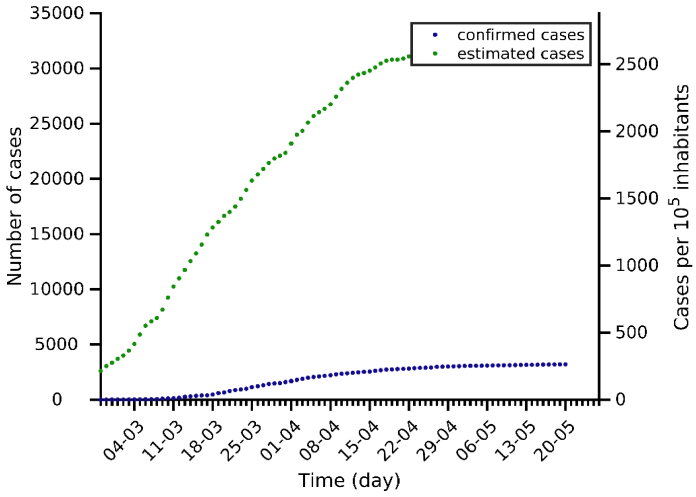


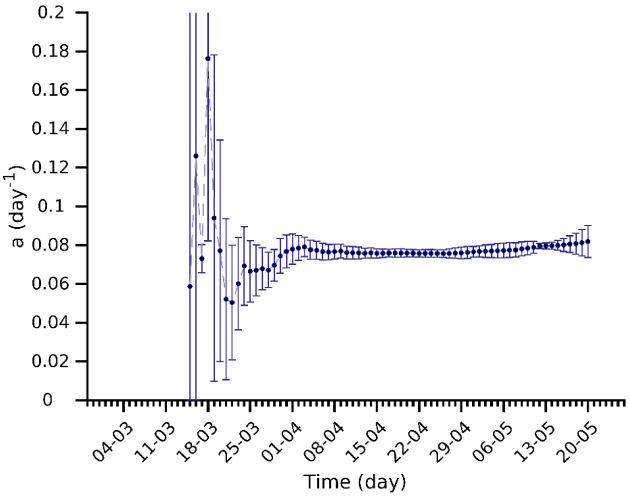


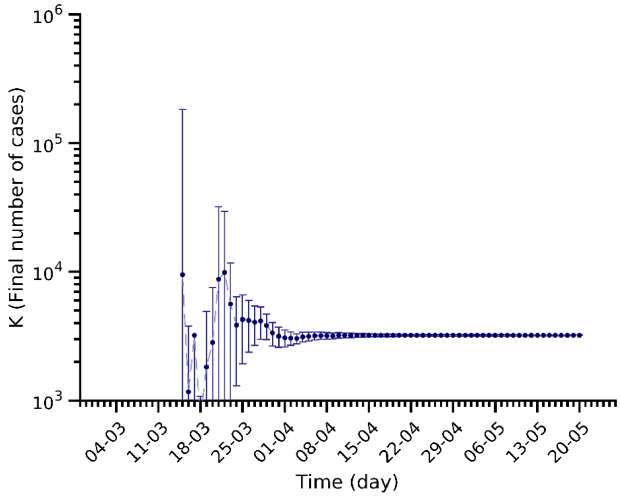


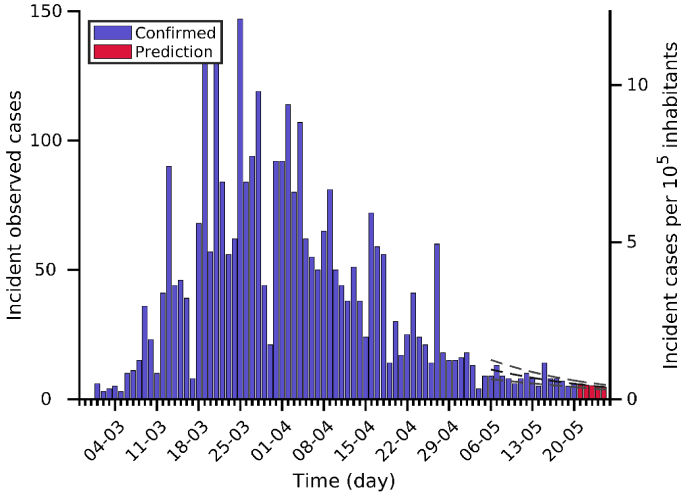
Friuli Venezia Giulia 20-05-2020. Population: 1.2M. Current cumulated incidence: 2

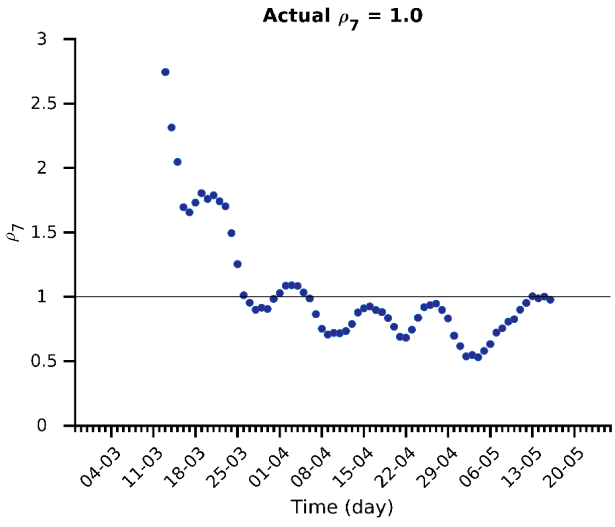


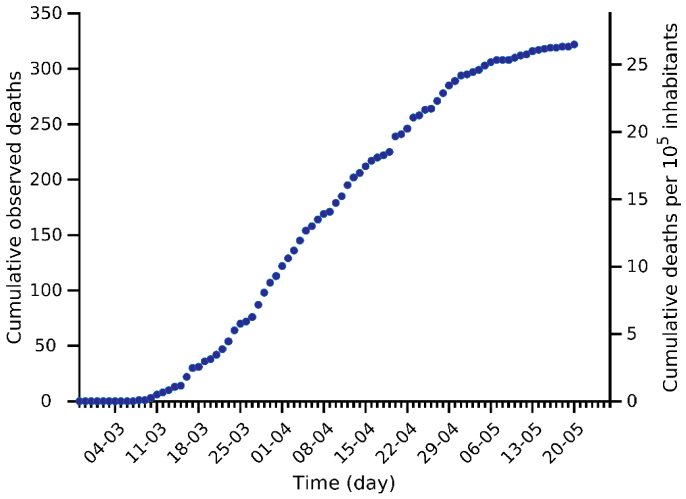


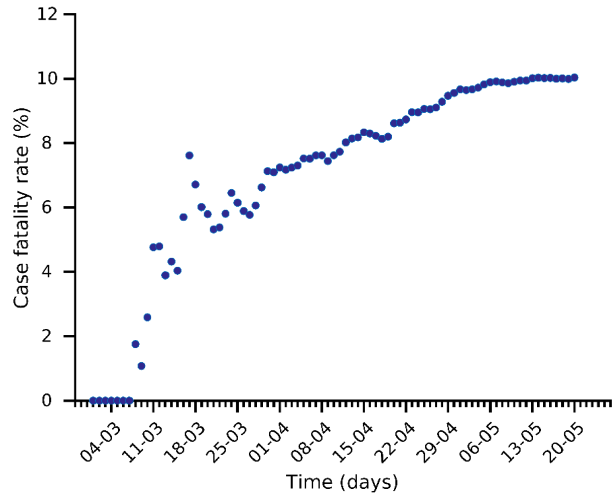




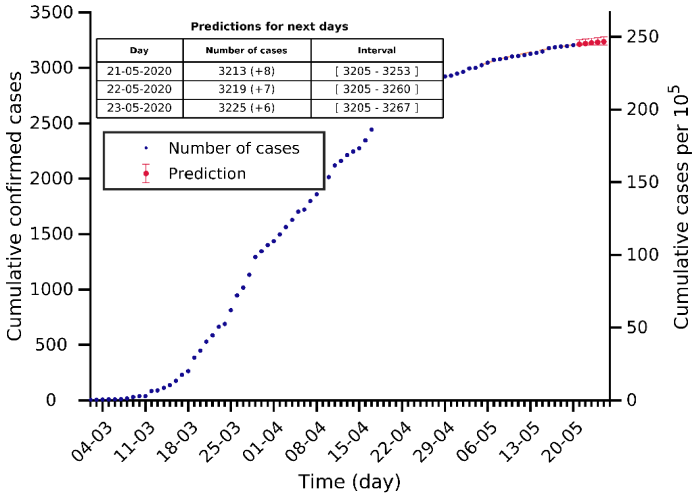


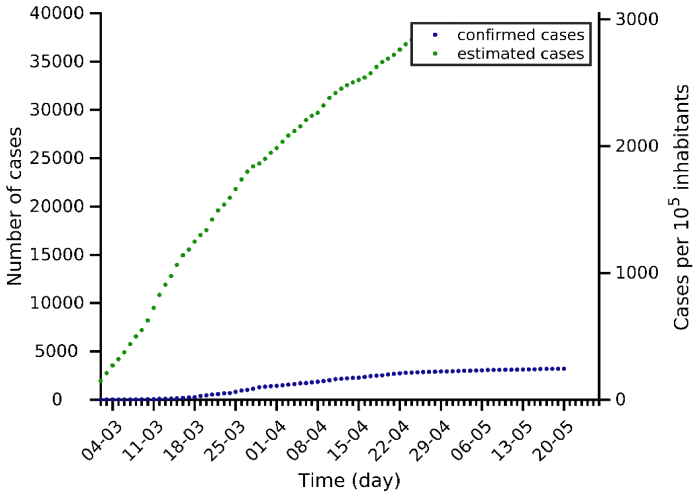


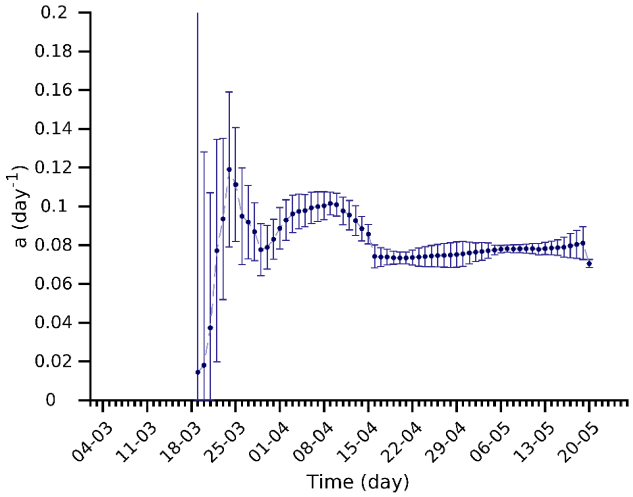


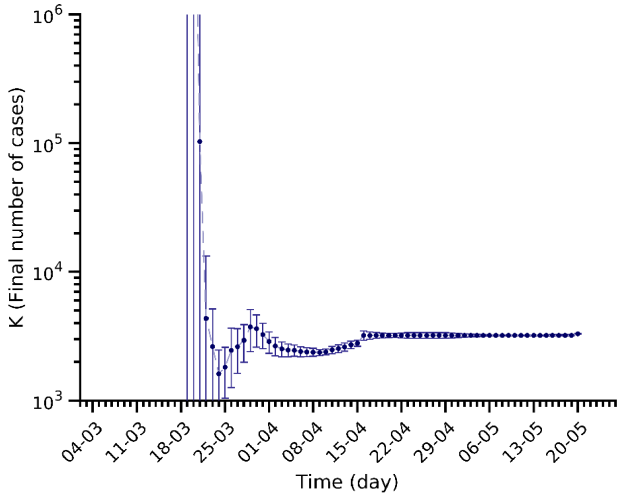


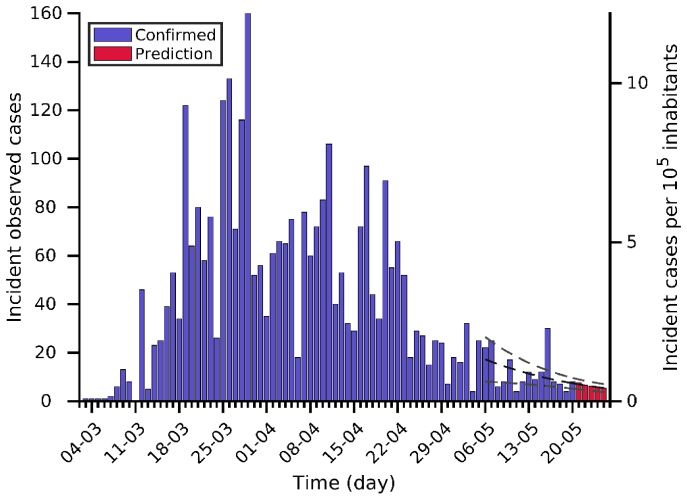
Abruzzo 20-05-2020. Population: 1.3M. Current cumulated incidence: 244/10^5

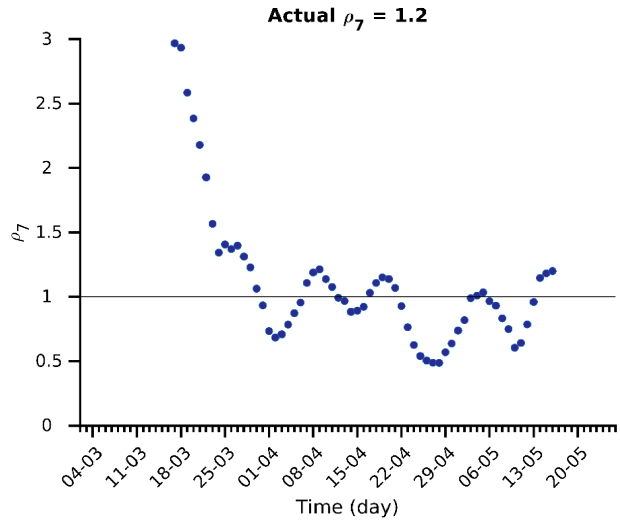


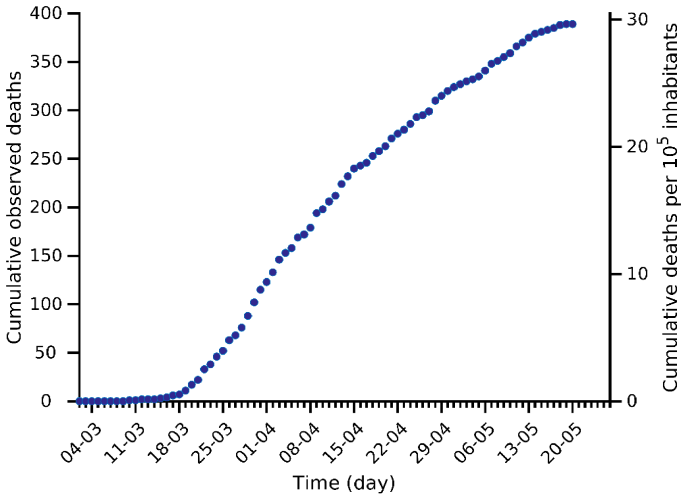


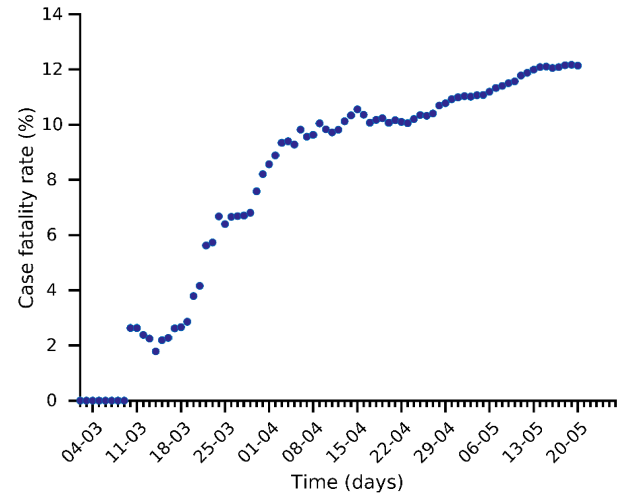




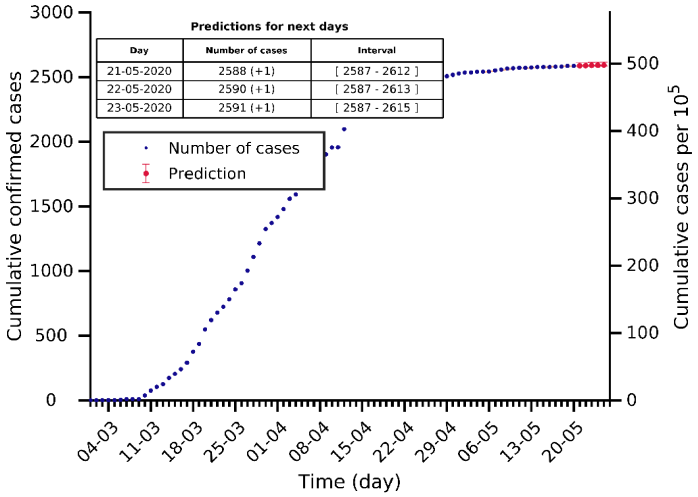


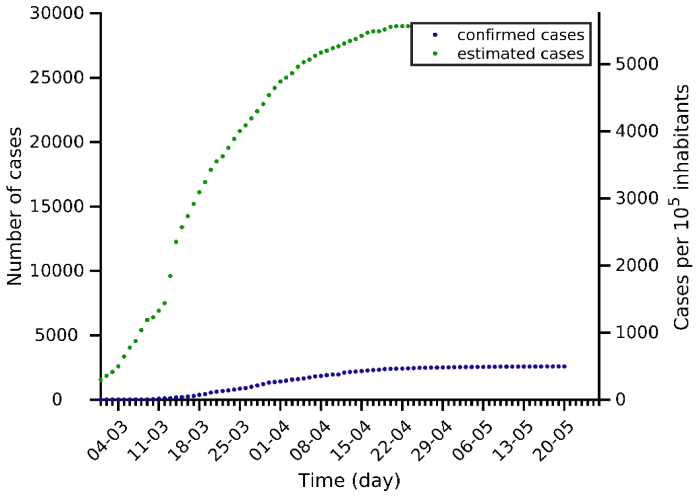


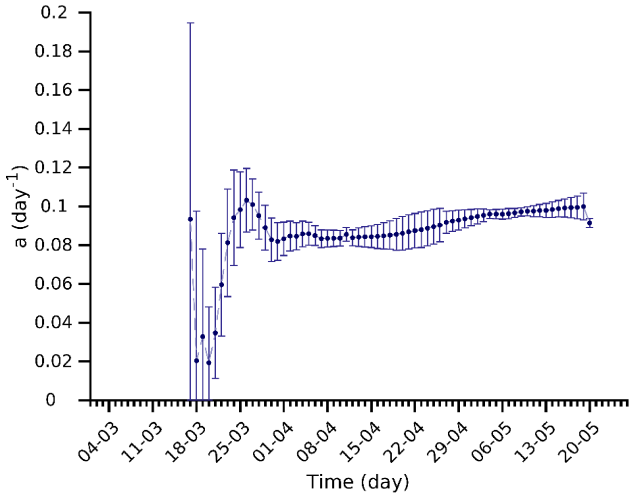


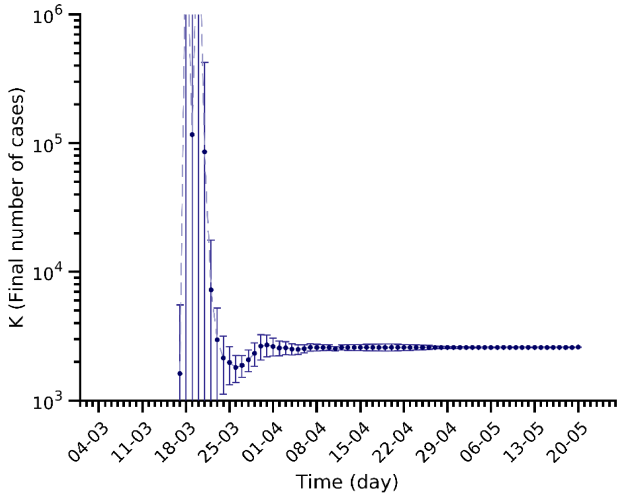


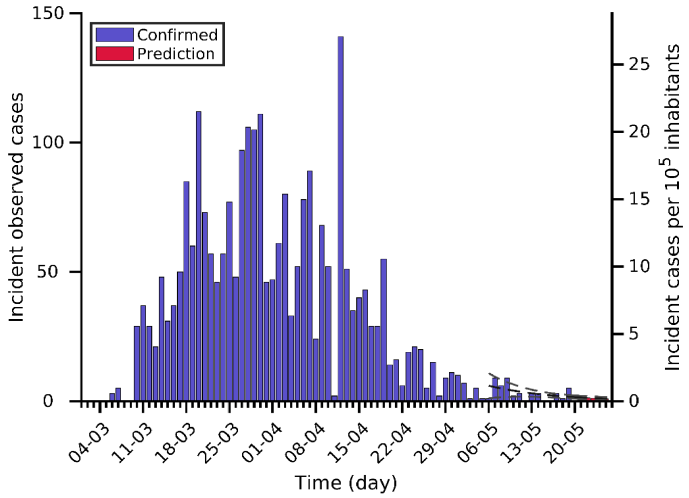
Bolzano 20-05-2020. Population: 0.5M. Current cumulated incidence: 497/10⁵

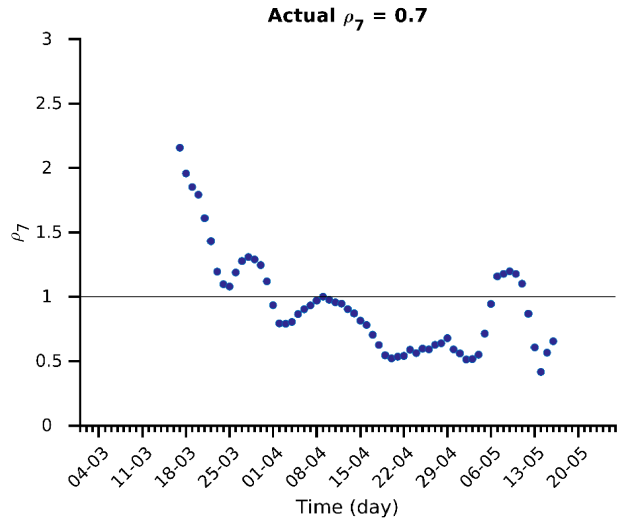


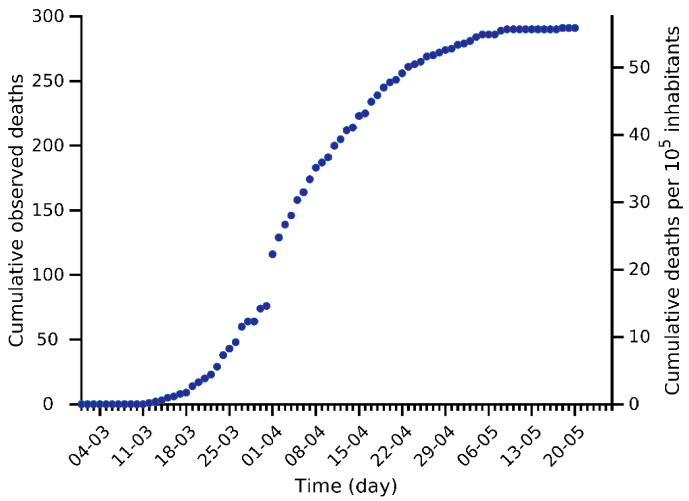


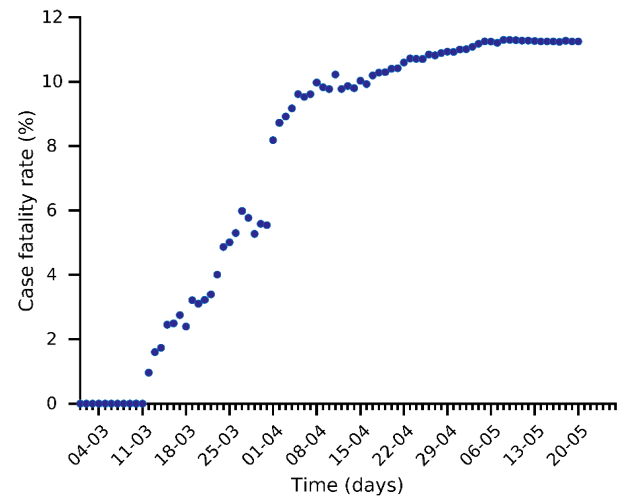




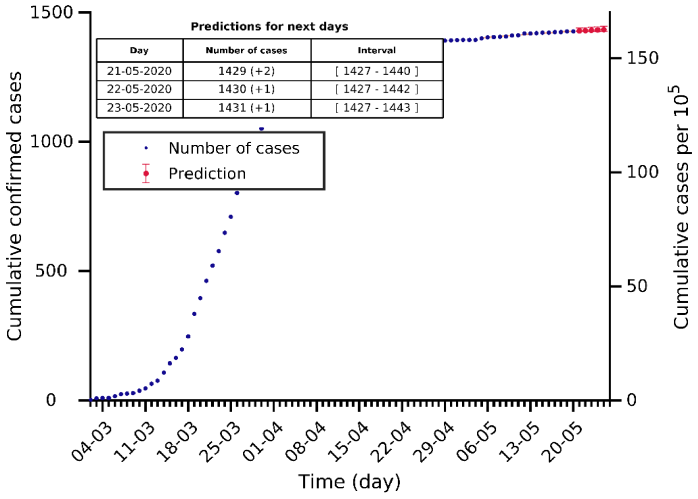


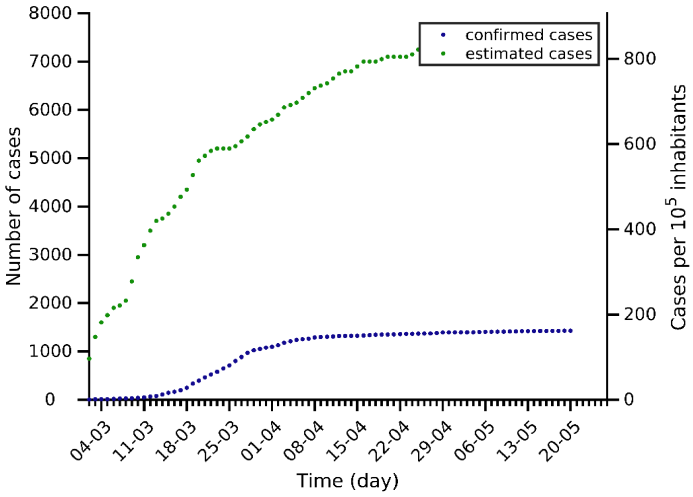


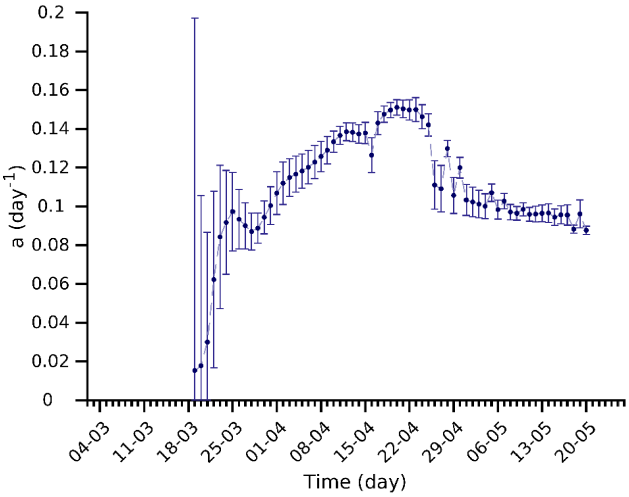


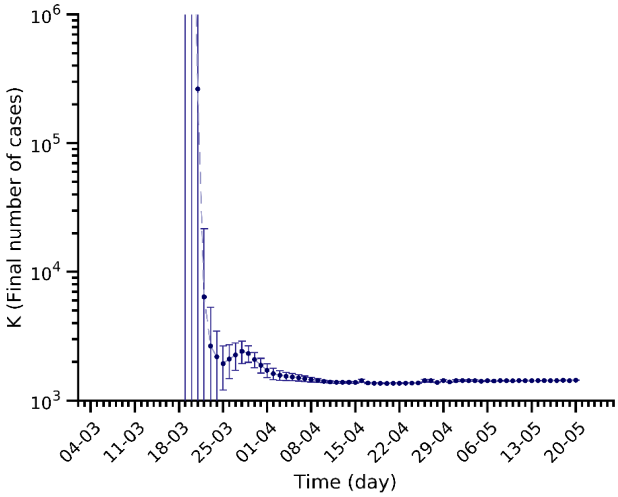


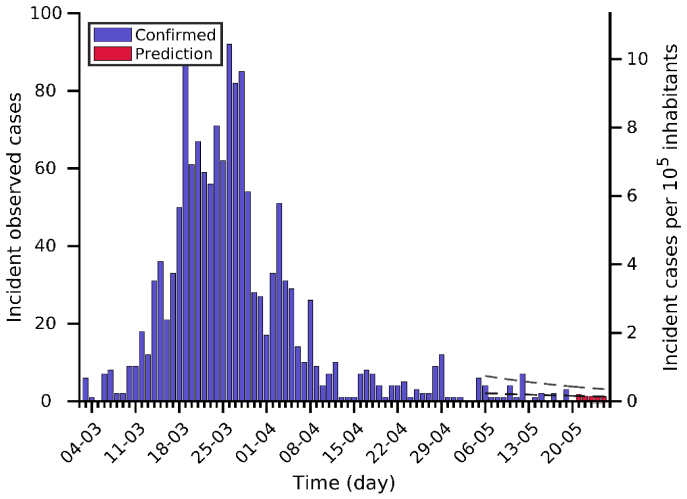
Umbria 20-05-2020. Population: 0.9M. Current cumulated incidence: $162/10^5$

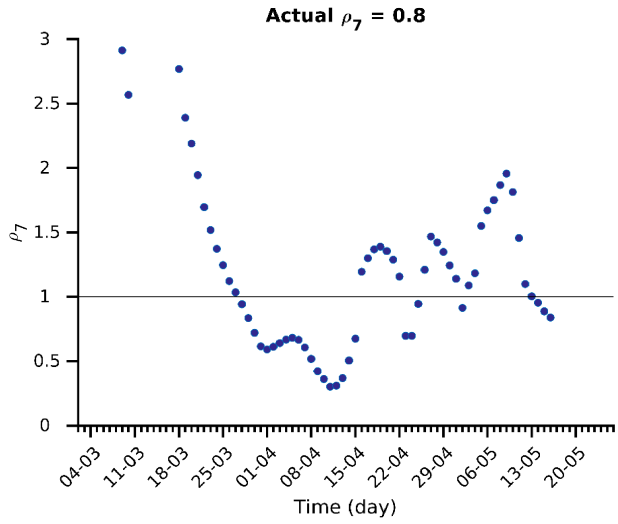


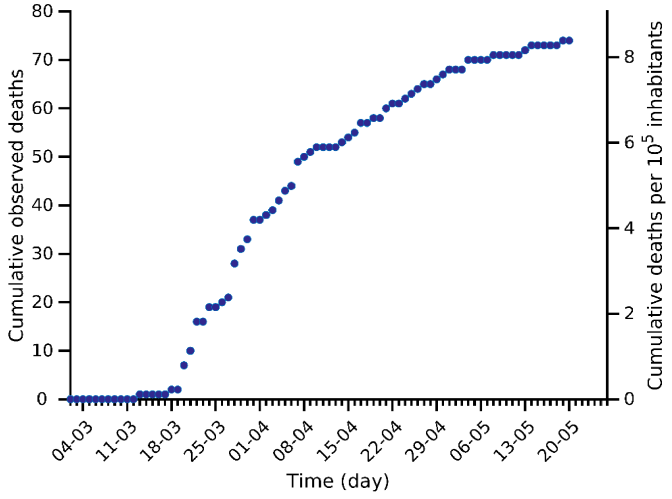


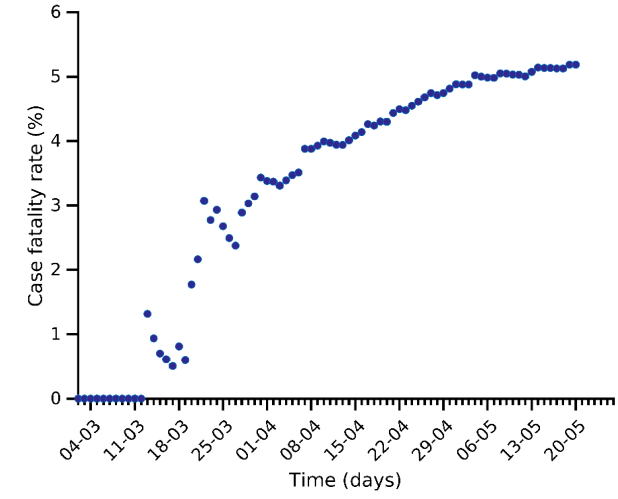




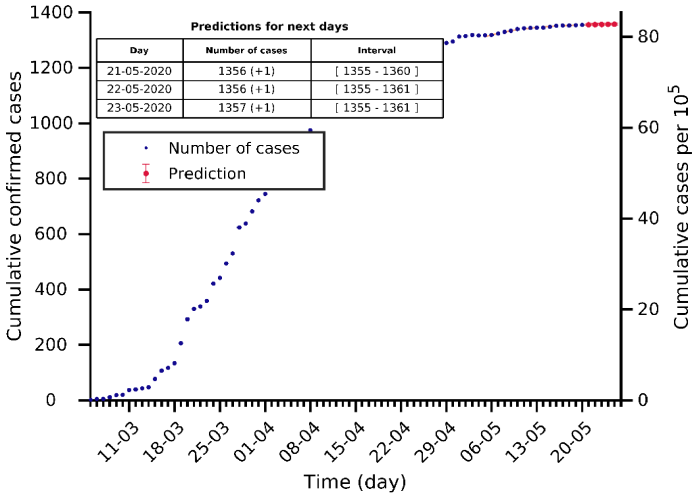


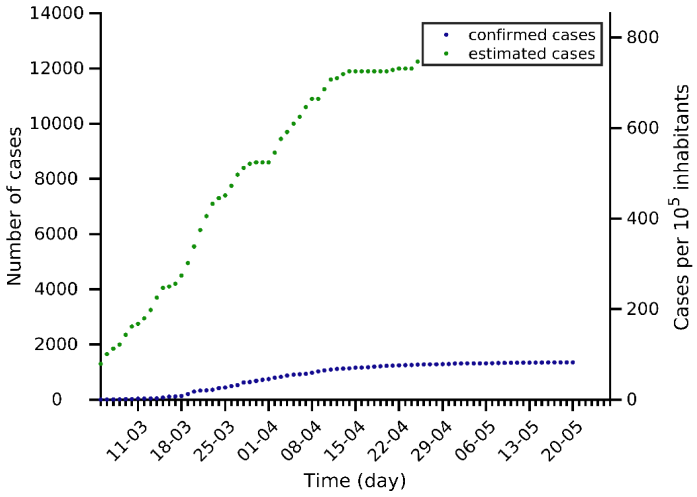


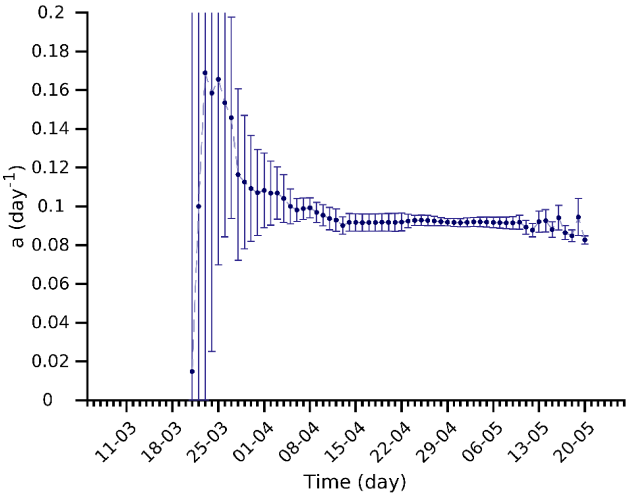


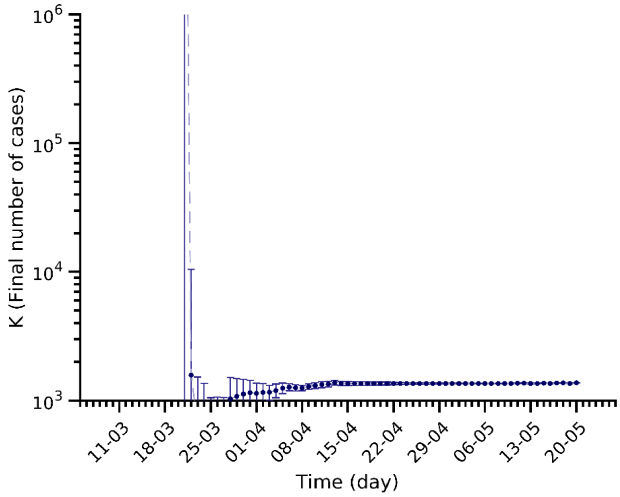


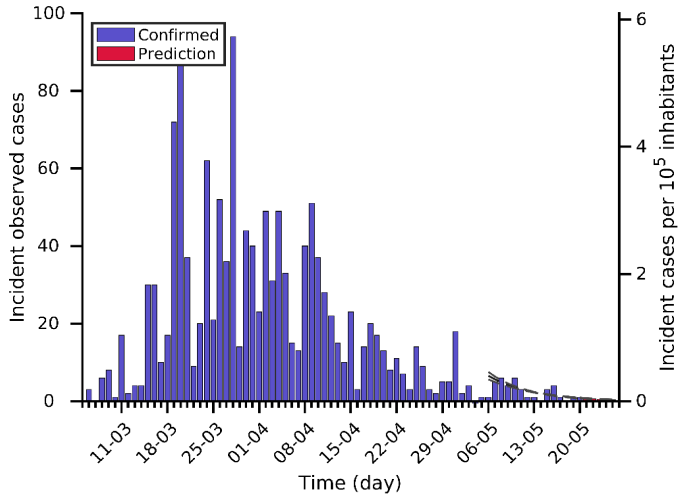
Sardegna 20-05-2020. Population: 1.6M. Current cumulated incidence: 83/10^5

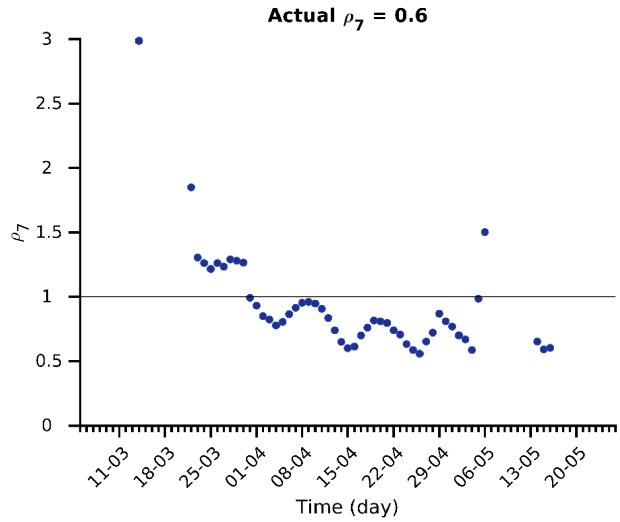


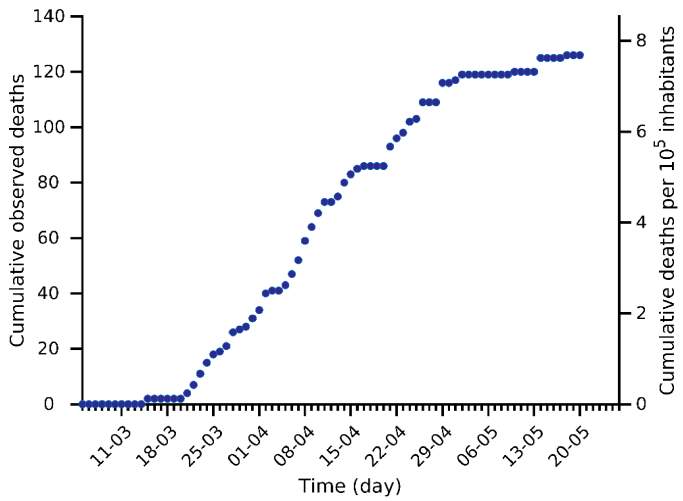


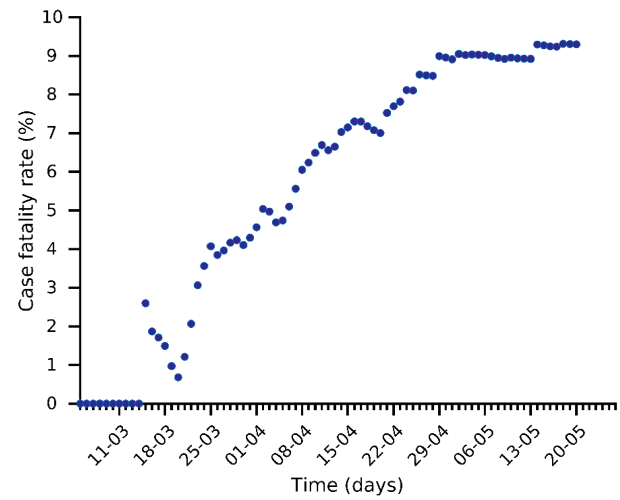




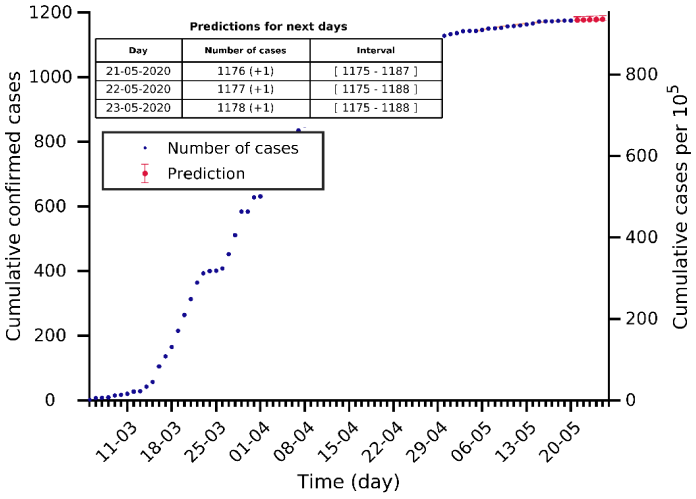


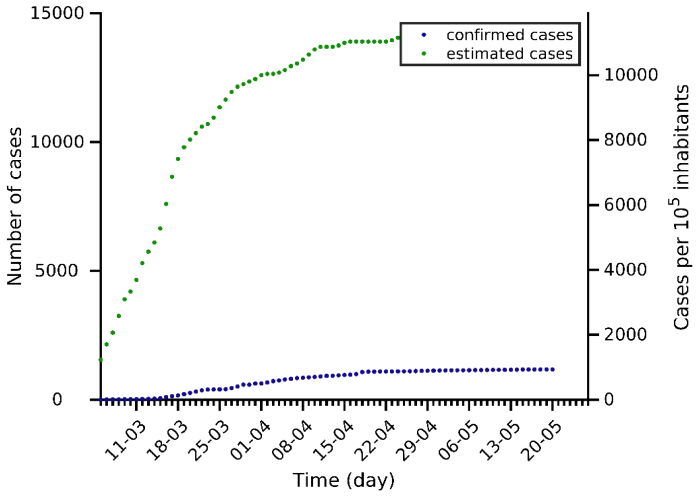


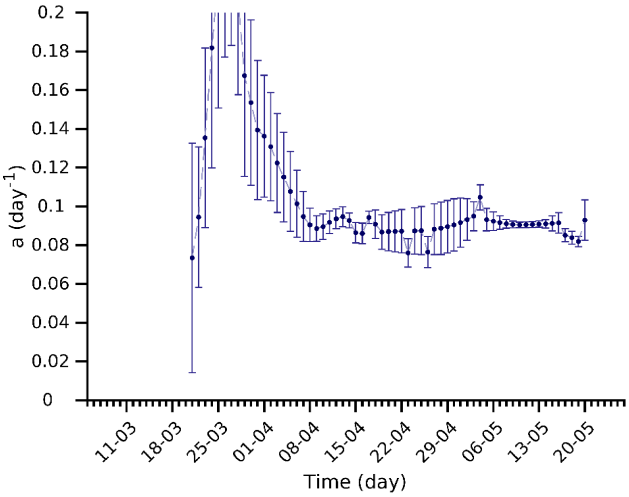


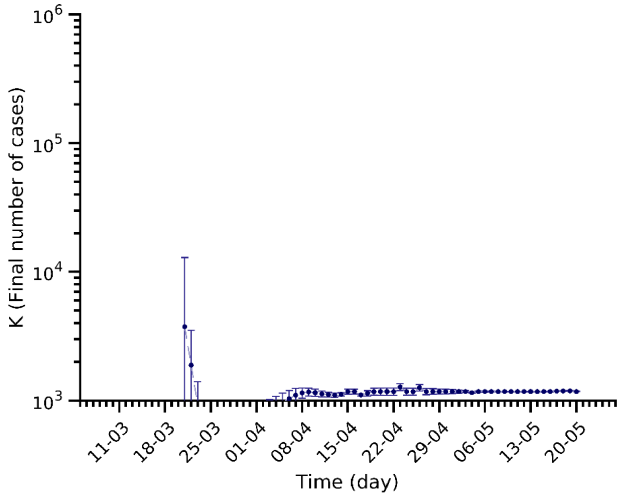


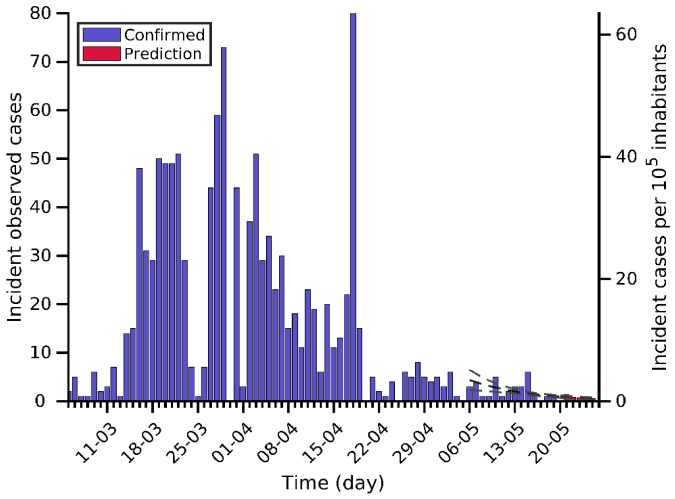
Valle dAosta 20-05-2020. Population: 0.1M. Current cumulated incidence: 933/10^5

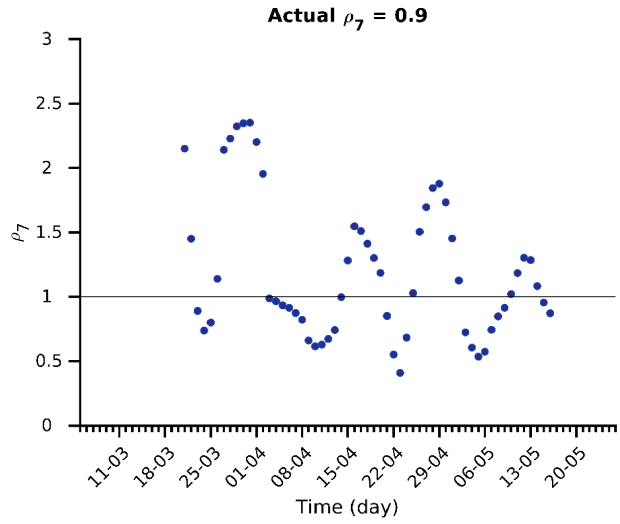


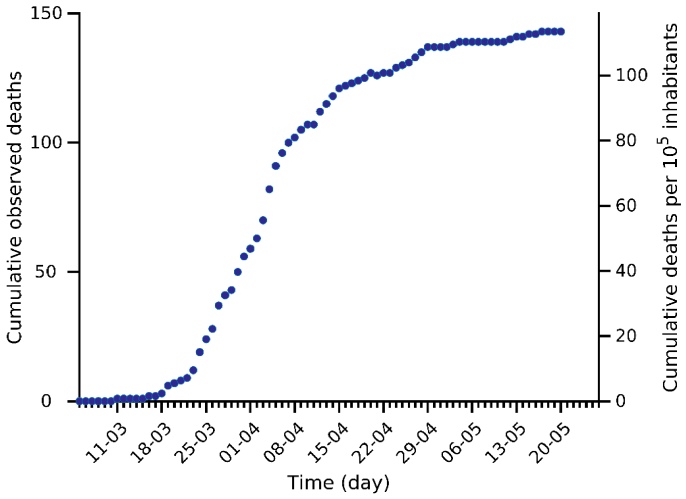


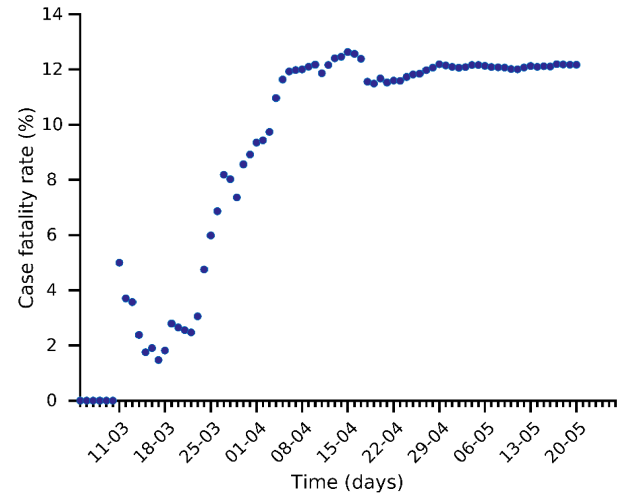




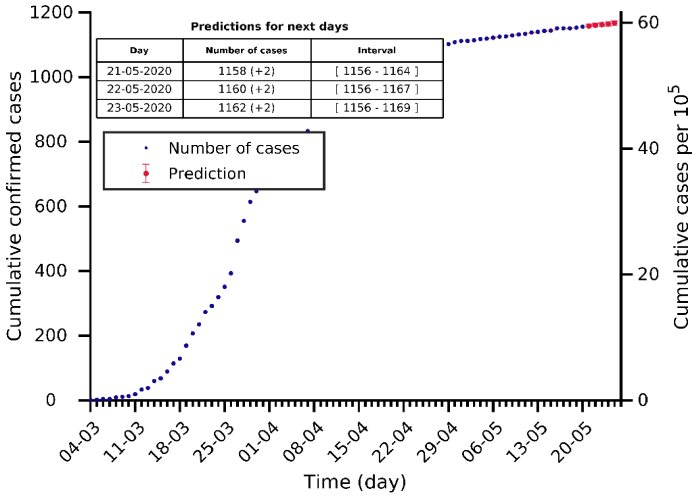


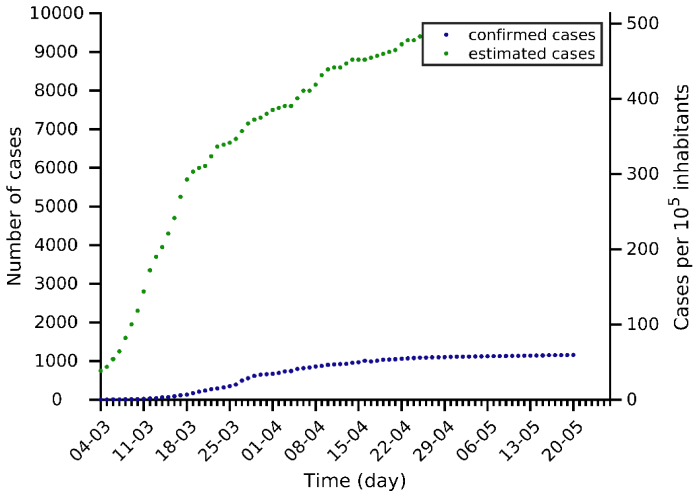


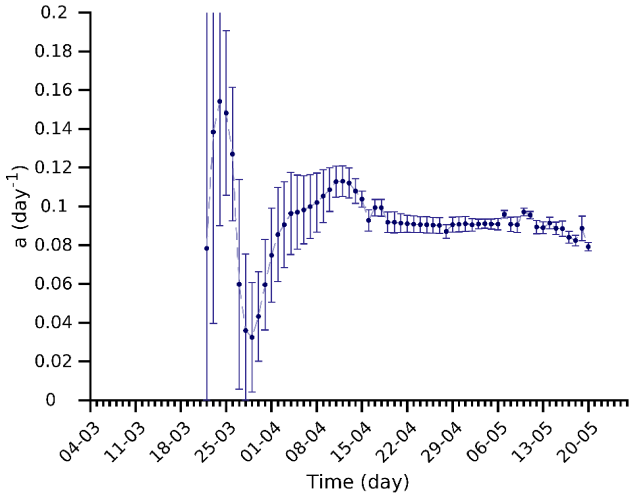


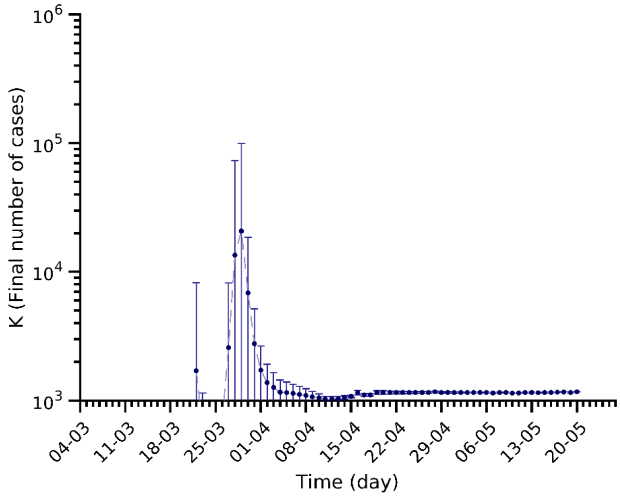


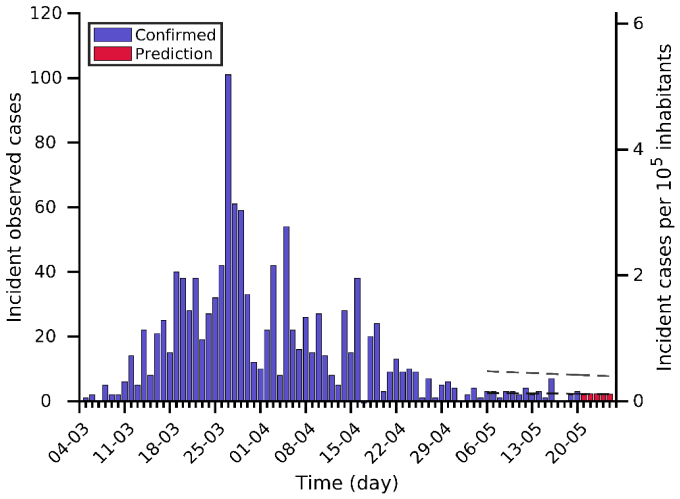
Calabria 20-05-2020. Population: 1.9M. Current cumulated incidence: $59/10^5$

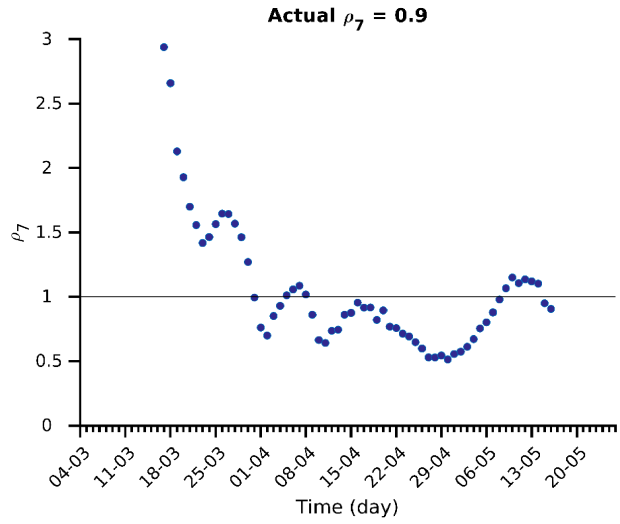


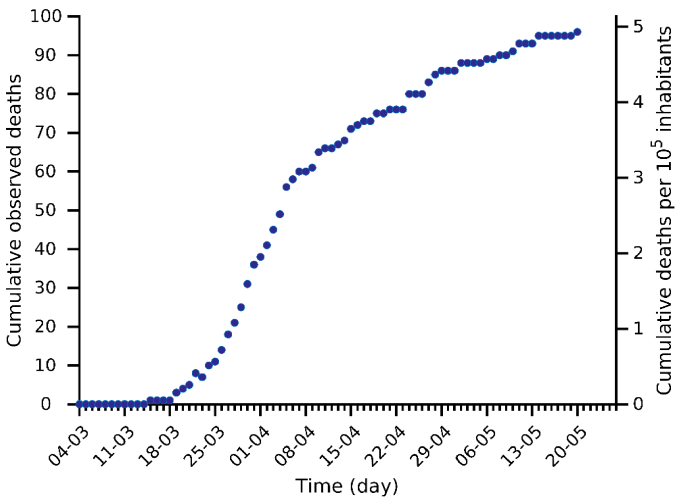


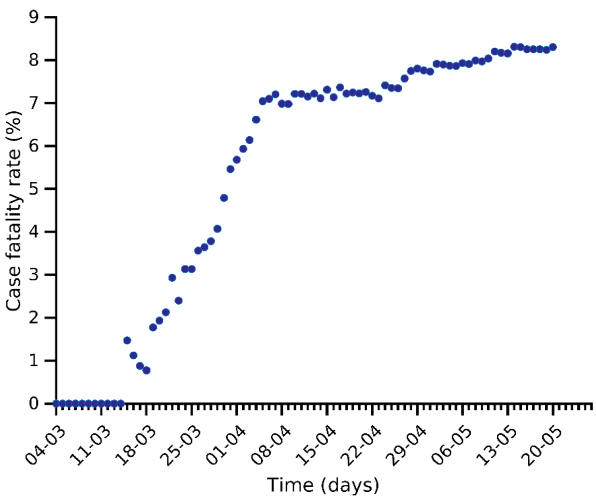




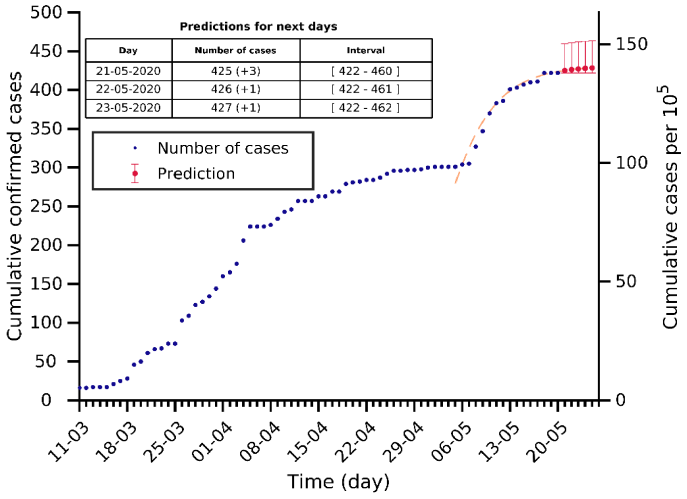


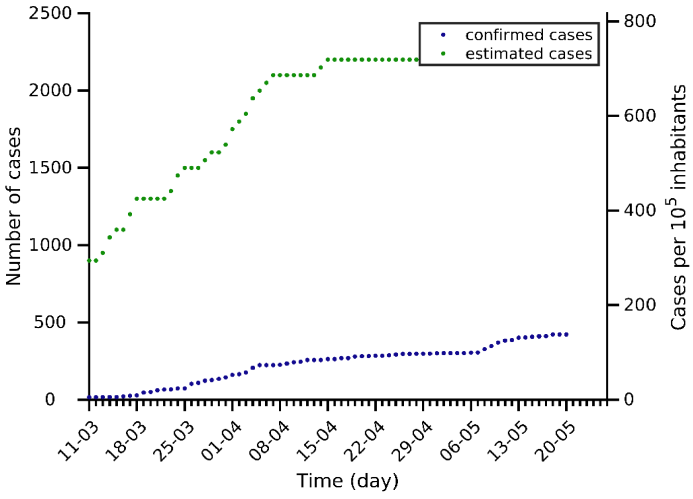


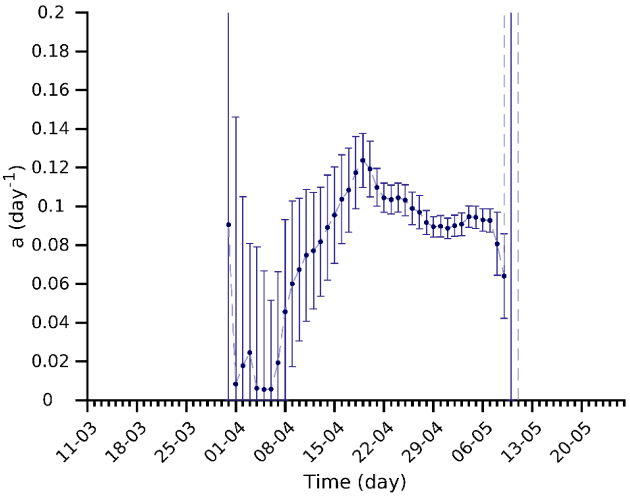


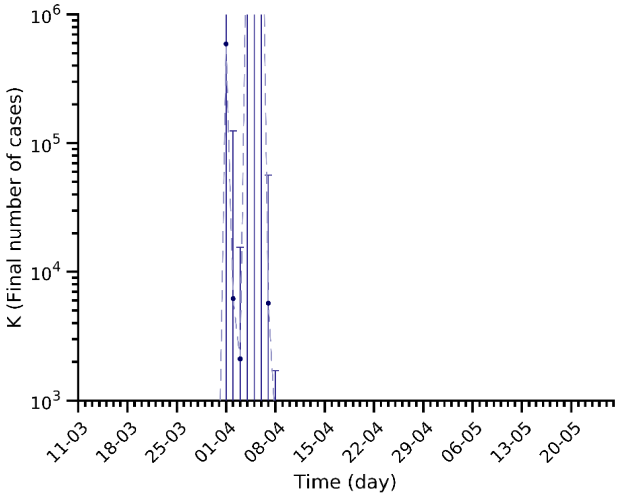


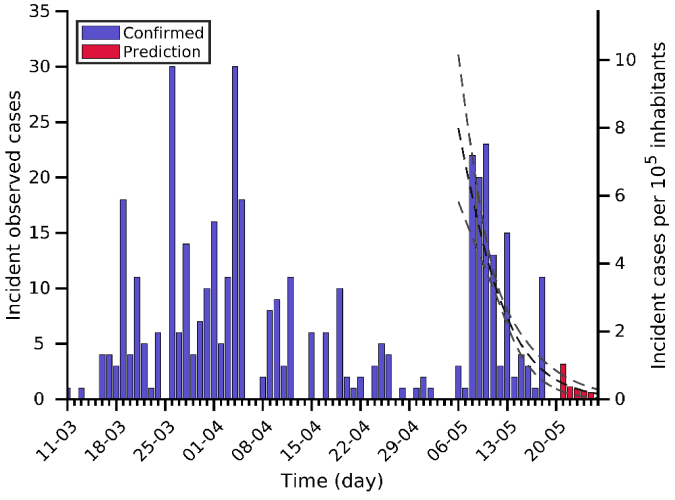
Molise 20-05-2020. Population: 0.3M. Current cumulated incidence: 138/10^5

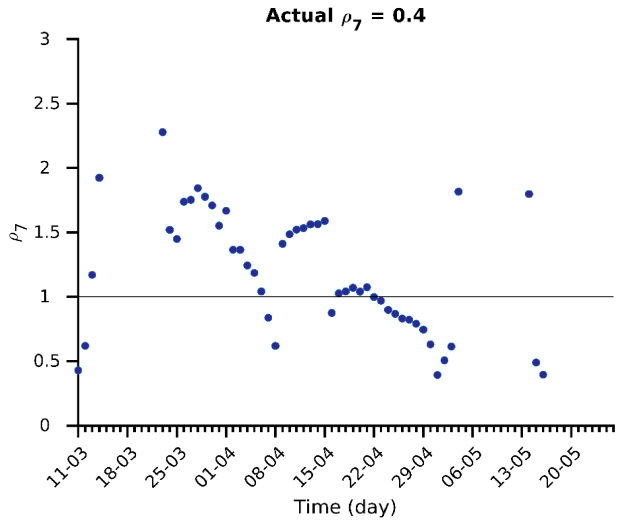


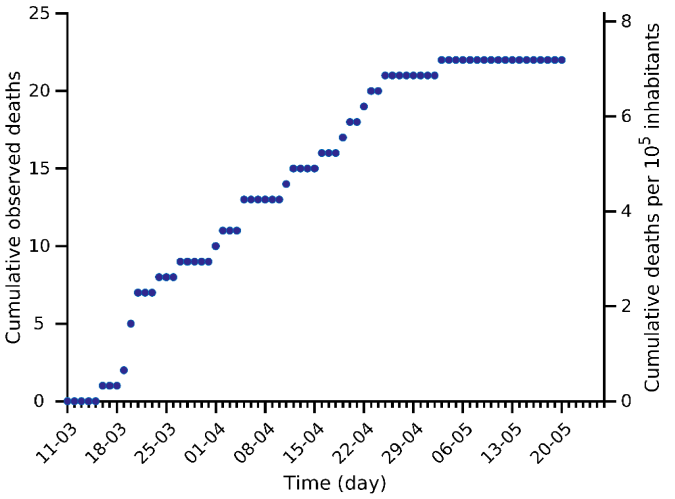


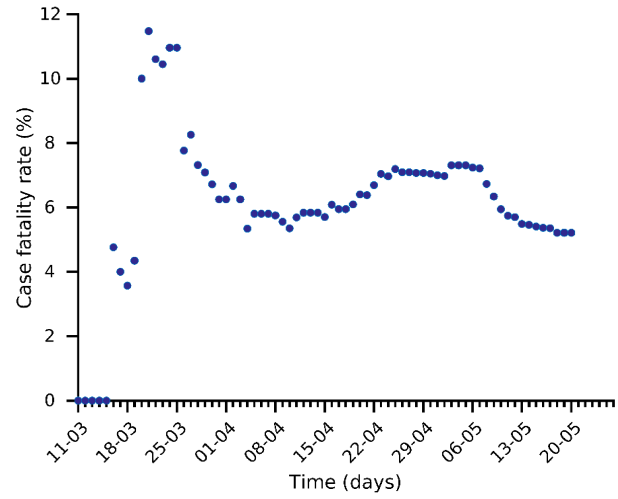




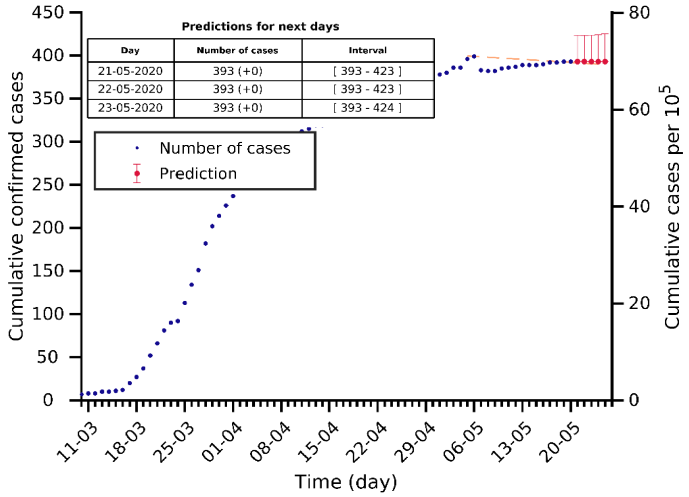


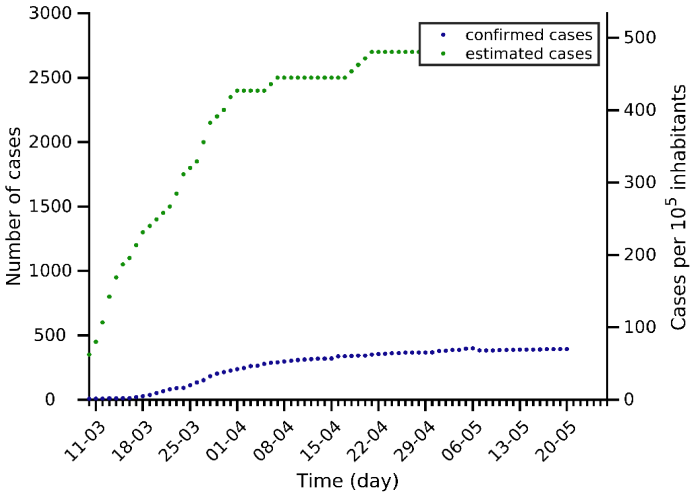


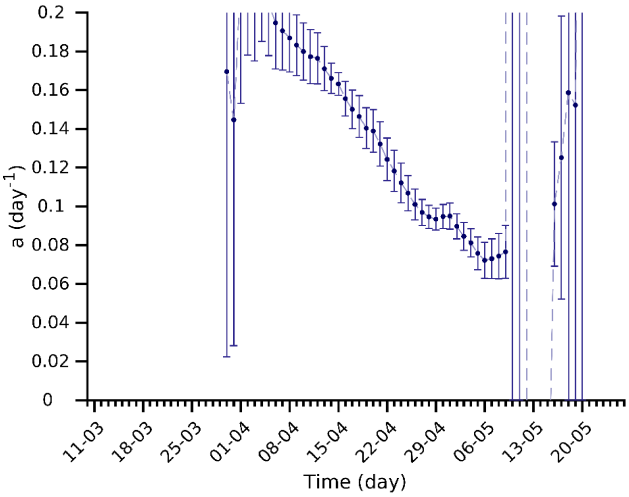


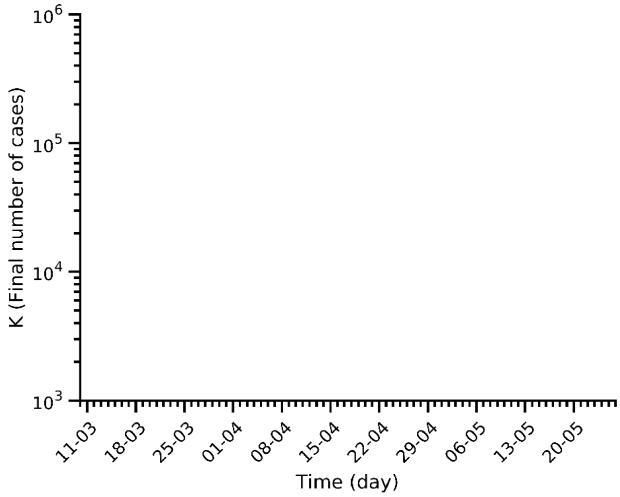


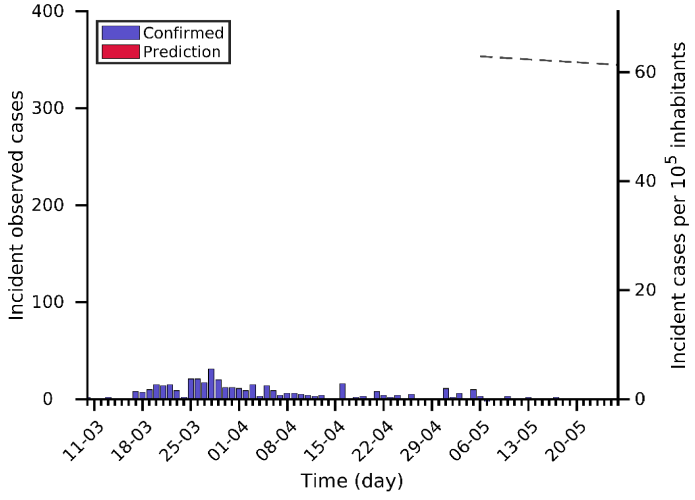
Basilicata 20-05-2020. Population: 0.6M. Current cumulated incidence: 70/10^5

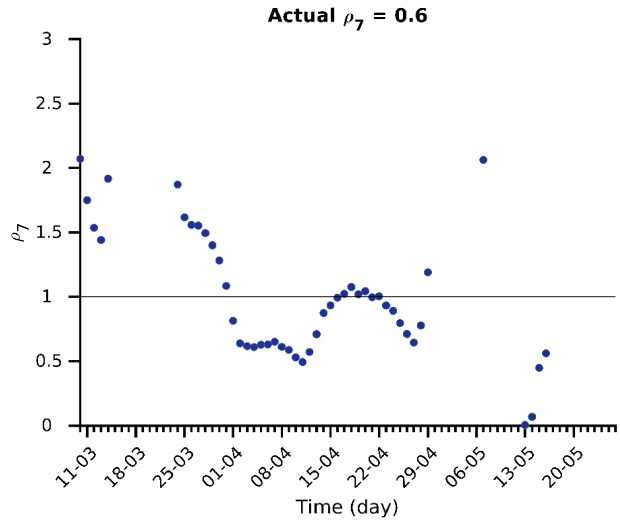


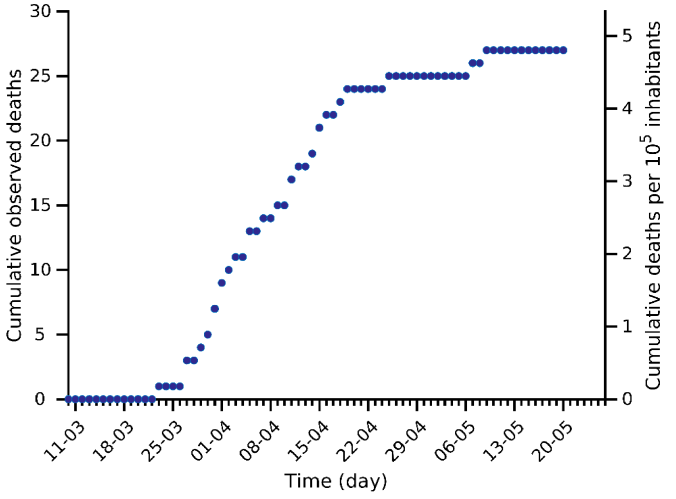


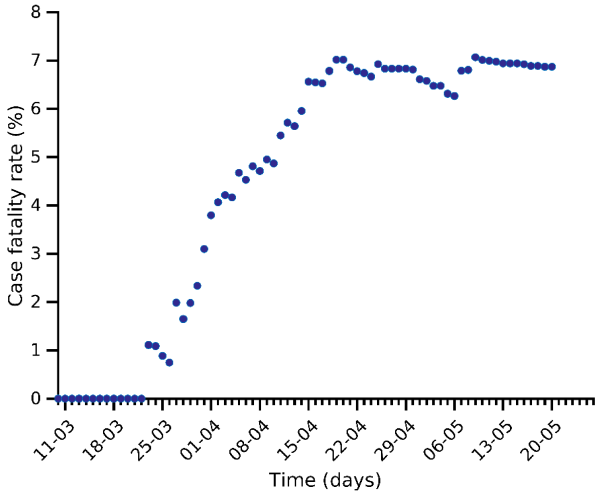












Methods

Methods

(1) Data source

Data are daily obtained from World Health Organization (WHO) surveillance reports[3], from European Centre for Disease Prevention and Control (ECDC)[4] and from Ministerio de Sanidad[5]. These reports are converted into text files that can be processed for subsequent analysis. Daily data comprise, among others: total confirmed cases, total confirmed new cases, total deaths, total new deaths. It must be considered that the report is always providing data from previous day. In the document we use the date at which the datapoint is assumed to belong, i.e., report from 15/03/2020 is giving data from 14/03/2020, the latter being used in the subsequent analysis.

(2) Data processing and plotting

Data are initially processed with Matlab in order to update timeseries, i.e., last datapoints are added to historical sequences. These timeseries are plotted for EU individual countries and for the UE as a whole:

- ✓ Number of cumulated confirmed cases, in blue dots
- ✓ Number of reported new cases
- ✓ Number of cumulated deaths

Then, two indicators are calculated and plotted, too:

- ✓ Number of cumulated deaths divided by the number of cumulated confirmed cases, and reported as a percentage; it is an indirect indicator of the diagnostic level.
- ✓ ρ: this variable is related with the reproduction number, i.e., with the number of new infections caused by a single case. It is evaluated as follows for the day before last report (*t-1*):

$$\rho(t-1) = \frac{N_{new}(t) + N_{new}(t-1) + N_{new}(t-2)}{N_{new}(t-5) + N_{new}(t-6) + N_{new}(t-7)}$$

where $N_{new}(t)$ is the number of new confirmed cases at day *t*.

(3) Classification of countries according to their status in the epidemic cycle

The evolution of confirmed cases shows a biphasic behaviour:

(I) an initial period where most of the cases are imported;
(II) a subsequent period where most of new cases occur because of local transmission.

Once in the stage II, mathematical models can be used to track evolutions and predict tendencies. Focusing on countries that are on stage II, we classify them in three groups:

- Group A: countries that have reported more than 100 cumulated cases for 10 consecutive days or more;
- Group B: countries that have reported more than 100 cumulated cases for 7 to 9 consecutive days;
- Group C: countries that have reported more than 100 cumulated cases for 4 to 6 days.

[3] https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports
[4] https://www.ecdc.europa.eu/en/geographical-distribution-2019-ncov-cases
[5] https://www.mscbs.gob.es/profesionales/saludPublica/ccayes/alertasActual/nCov-China/situacionActual.htm https://github.com/datadista/datasets/tree/master/COVID%2019 , https://covid19.isciii.es/

(4) Fitting a mathematical model to data

Previous studies have shown that Gompertz model[6] correctly describes the Covid-19 epidemic in all analysed countries. It is an empirical model that starts with an exponential growth but that gradually decreases its specific growth rate. Therefore, it is adequate for describing an epidemic that is characterized by an initial exponential growth but a progressive decrease in spreading velocity provided that appropriate control measures are applied.

Gompertz model is described by the equation:

$$N(t) = K\, e^{-ln\left(\frac{K}{N_0}\right)\cdot e^{-a\cdot(t-t_0)}}$$

where *N(t)* is the cumulated number of confirmed cases at *t* (in days), and N_0 is the number of cumulated cases the day at day t_0. The model has two parameters:

- ✓ *a* is the velocity at which specific spreading rate is slowing down;
- ✓ *K* is the expected final number of cumulated cases at the end of the epidemic.

This model is fitted to reported cumulated cases of the UE and of countries in stage II that accomplish two criteria: 4 or more consecutive days with more than 100 cumulated cases, and at least one datapoint over 200 cases. Day t_0 is chosen as that one at which *N(t)* overpasses 100 cases. If more than 15 datapoints that accomplish the stated criteria are available, only the last 15 points are used. The fitting is done using Matlab's Curve Fitting package with Nonlinear Least Squares method, which also provides confidence intervals of fitted parameters (*a* and *K*) and the R^2 of the fitting. At the initial stages the dynamics is exponential and *K* cannot be correctly evaluated. In fact, at this stage the most relevant parameter is *a*. Fitted curves are incorporated to plots of cumulative reported cases with a dashed line. Once a new fitting is done, two plots are added to the country report:

- ✓ Evolution of fitted *a* with its error bars, i.e., values obtained on the fitting each day that the analysis has been carried out;
- ✓ Evolution of fitted *K* with its error bars, i.e., values obtained on the fitting each day that the analysis has been carried out; if lower error bar indicates a value that is lower than current number of cases, the error bar is truncated.

These plots illustrate the increase in fittings' confidence, as fitted values progressively stabilize around a certain value and error bars get smaller when the number of datapoints increases. In fact, in the case of countries, they are discarded and set as "Not enough data" if $a>0.2\ day^{-1}$, if $K>10^6$ or if the error in K overpasses 10^6.

It is worth to mention that the simplicity of this model and the lack of previous assumptions about the Covid-19 behaviour make it appropriate for universal use, i.e., it can be fitted to any country independently of its socioeconomic context and control strategy. Then, the model is capable of quantifying the observed dynamics in an objective and standard manner and predicting short-term tendencies.

(5) Using the model for predicting short-term tendencies

The model is finally used for a short-term prediction of the evolution of the cumulated number of cases. The predictions increase their reliability with the number of datapoints used in the fitting. Therefore, we consider three levels of prediction, depending on the country:

[6] Madden LV. Quantification of disease progression. Protection Ecology 1980; **2**: 159-176.

- Group A: prediction of expected cumulated cases for the following 3-5 days[7];
- Group B: prediction of expected cumulated cases for the following 2 days;
- Group C: prediction of expected cumulated cases for the following day.

The confidence interval of predictions is assessed with the Matlab function `predint`, with a 99% confidence level. These predictions are shown in the plots as red dots with corresponding error bars, and also gathered in the attached table. For series longer than 9 timepoints, last 3 points are weighted in the fitting so that changes in tendencies are well captured by the model.

(6) Estimating non-diagnosed cases

Lethality of Covid-19 has been estimated at around 1 % for Republic of Korea and the Diamond Princess cruise. Besides, median duration of viral shedding after Covid-19 onset has been estimated at 18.5 days for non-survivors[8] in a retrospective study in Wuhan. These data allow for an estimation of total number of cases, considering that the number of deaths at certain moment should be about 1 % of total cases 18.5 days before. This is valid for estimating cases of countries at stage II, since in stage I the deaths would be mostly due to the incidence at the country from which they were imported. We establish a threshold of 50 reported cases before starting this estimation.

Reported deaths are passed through a moving average filter of 5 points in order to smooth tendencies. Then, the corresponding number of cases is found assuming the 1 % lethality. Finally, these cases are distributed between 18 and 19 days before each one.

[7] At this moment we are testing predictions at 4 days for countries with more than 100 cumulated cases for 13-15 consecutive days, and 5 days for 16 or more days.

[8] Zhou et al., 2020. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. The Lancet; March 9, doi: 10.1016/S0140-6736(20)30566-3